1998

Involvement of Human Immunodeficiency Virus Type-1 Splice Sites in the Cytoplasmic Accumulation of Viral RNA

Keith Thomas Borg
Medical University of South Carolina

Involvement of Human Immunodeficiency Virus Type-1 Splice Sites in the Cytoplasmic Accumulation of Viral RNA

Keith Thomas Borg

A dissertation submitted to the faculty of the
Medical Unversity of South Carolina in partial fulfillment for the degree of
Doctor of Philosophy in Microbiology and Immunology
in the College of Graduate Studies

Department Microbiology and Immunology.

April 24, 1998

approved by

Chairman, Advisory Committee
Salvatore Arrigo, Ph.D.

Michael Schmidt, Ph.D.

Joseph Dolan, Ph.D.

James Norris, Ph.D.

Edwin Brown, M.D.

E. Carwile LeRoy, M.D.

This work is dedicated to my parents-

Thomas Keith and Karen Lynn Borg

They instilled in me not only a love of science but also a pattern of observation and questioning the world at a young age.

For their efforts I will always be indebted.

Acknowledgements

This work could not have been accomplished without very significant contributions from a large number of people. I wish to thank my committee- Salvatore Arrigo, Michael Schmidt, Jim Norris, Ed Brown, Joe Dolan, and Carwile LeRoy. I would also like to thank Gill Galbraith who graciously offered to serve as a proxy committee member. Perry Halushka has, in addition to running a strong MSTP program, been extremely helpful in providing both advice and resources for the completion of this work. Thank you also to Jim Norris and Carwile Leroy who made sure that there would be space and funding for the completion of these studies. I would also like to thank Jeff Patton for his helpful advice and support.

Special thanks to all my fellow students Lia Campbell, Justin Favaro, Joel Gatlin, and Kim Huffman for making the journey not only more enjoyable but also making it possible.

Without the extraordinary support and encouragement of Kate Hendrickson-Borg none of this work could have been completed.

And thanks to Joe Dolan for "how about 8/23?" without which I might never have seen the light.

I would like to thank my advisor and mentor Sal Arrigo who has been instrumental not only in the development of this work but also in teaching me how to be a scientist.

Without the selfless efforts of Mike Schmidt this thesis would not and could not have been finished, thank you for your help.

I would also like to thank the boys- stan, eric, kyle and kenny for inspiration and perspective. Also special gratitude to Jake, who did not eat the thesis but has seemingly eaten everything else.

To quote from David Lindberg in his preface to *The Beginnings of Western Science* "If this list is remarkable for its length, I can only explain that I needed all the help I could get."

Table of Contents

List of Figures

Abstract

KEITH BORG. Involvement of Human Immunodeficiency Virus Type-1 Splice Sites in the Cytoplasmic Accumulation of Viral RNA. (Under the direction of SALVATORE ARRIGO).

The Human Immunodeficiency Virus Type 1 is a complex retrovirus, which produces a large number of proteins from a 10 Kb genome. The virus utilizes three overlapping open reading frames, a ribosomal frameshift event and three different classes of RNA (unspliced, partially spliced and fully spliced RNA) to perform this feat. To regulate the expression of these RNAs, HIV-1 requires Rev, a viral *trans* regulatory protein. The fully spliced RNAs are not dependent upon Rev for their expression. The unspliced and partially spliced RNAs are dependent on Rev to produce proteins. Rev binds to the Rev Response Element (RRE) located in the Rev dependent RNAs. The sequences responsible for the lack of production of protein in the absence of Rev, have been termed Cis-acting Repressor Sequences (CRS). Little is known about the mechanism or nature of CRS. It is believed that CRS serve to confer Rev dependence on the RNA and allow for the regulation of the RNA at a number of post-transcriptional levels. It is our hypothesis that, since Rev regulates the production of the different classes of mRNA, splice sites are intimately involved in the regulation of these RNAs by Rev. To define the role of Human Immunodeficiency Virus type 1 splice sites in the cytoplasmic accumulation of viral RNAs, sequential deletion mutagenesis on an infectious proviral clone of HIV-1 was performed. Deletion of the majority of intron sequences, containing previously identified CRS, did not attenuate CRS activity. Retention of either the first or second *tat* intron

preserved CRS activity. RNAs containing splice donor sequences, in the absence of known downstream splice acceptor sequences, retained CRS activity. These results indicate that the major HIV-1 splice donors can function as CRS and function to negatively regulate the cytoplasmic accumulation of HIV-1 RNAs in COS cells. In order to further analyze the role of a splice donor in Rev dependence, the wild type 5' splice donor of HIV-1 was mutated in the context of other *gag* sequences. The unspliced RNA produced by the mutant construct still required Rev for the cytoplasmic accumulation of the RNA. Despite deletion of the wild type 5' splice donor, splicing still occurred, as measured by usage of the *tat* splice acceptor. A cryptic splice donor was identified by PCR and subsequent cloning of the spliced RNA. The cryptic site is 5/9 to the consensus sequence and located immediately downstream of the initiation codon (ATG) for Gag. Analysis of RNA containing the cryptic splice donor demonstrated that Rev is still required for the cytoplasmic accumulation of the unspliced RNA, while spliced RNA is Rev independent. These results indicate that a cryptic splice donor can be activated when the wild-type splice donor is inactivated and that the cryptic splice donor retains Rev regulation. The findings also suggest the potential for cryptic splice sites to serve as CRS in determining the Rev-dependence of viral RNAs. These studies have investigated the critical role that splice sites can play in viral gene regulation and shown that splice sites, both wild-type and cryptic can function as CRS to control viral RNA expression.

Chapter 1

General Introduction

The complex gene regulation mechanisms of cells are utilized by viruses to regulate and facilitate programs of viral gene expression. To understand the coordinating mechanisms of viral gene regulation, it is crucial to understand the cellular processes involved in RNA processing in concert with gene expression. Eukaryotic processing of mRNA is a complex process that begins in the nucleus following transcription. An outline of the steps involved in RNA processing is depicted in Figure 1-1. Immediately following transcription, or even as it occurs, proteins called snRNPs (small nuclear ribonucleoproteins) associate with this primary transcript. snRNPs serve many functions during the processing of the transcript. The U1, U2, U5 and U4/6 snRNPs are involved in formation of the catalytic spliceosome which serves to splice the pre-mRNA (see figure 1.2) (Alberts *et al.*, 1994)

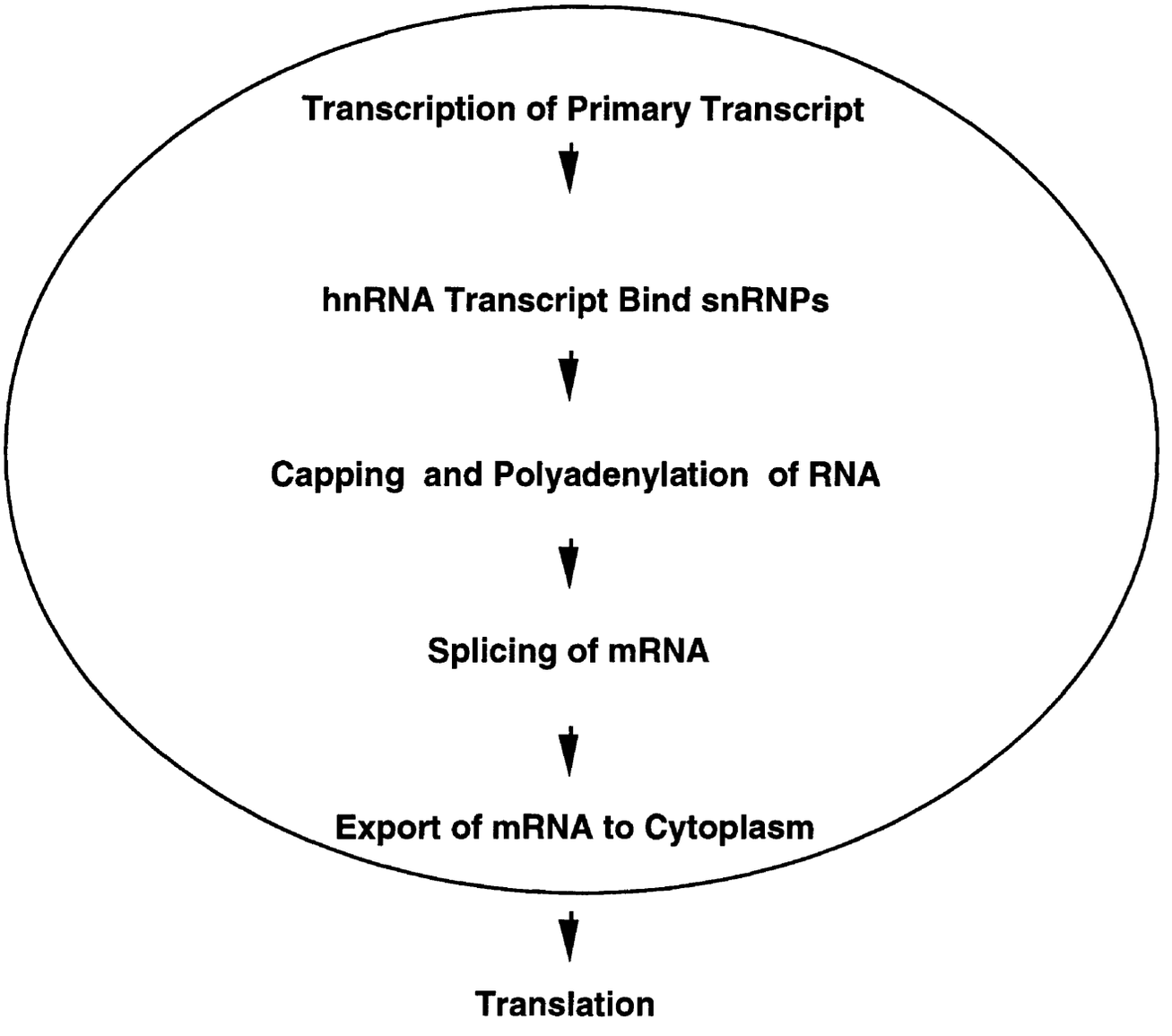


Figure 1.1 RNA Expression Steps involved in RNA processing in the nucleus and subsequent export from the nucleus to the cytoplasm.

After transcription and snRNP binding, the RNA is covalently modified at the 5' and 3' ends of the primary transcript. The 5' end is capped with a 7-methyl Guanosine residue; which is important not only in protecting the RNA from degradation, but also in the initiation of translation. The 3' end of the primary transcript is polyadenylated by poly-A polymerase which adds a series of adenosine residues about 20bp downstream of the polyadenylation signal- AAUAAA (Alberts *et al.*, 1994). This poly-A tail serves several functions in RNA processing including stabilization of the RNA and initiation of 60s ribosomal subunit binding for protein synthesis (Sachs and Wahle, 1993; Sachs, Bond, and Kornberg, 1986; Sachs and Davis, 1989; Sachs and Deardorff, 1992).

After initial processing the RNA is spliced via two trans-esterification reactions to produce a spliced product. This process takes place when snRNPs bound to the RNA catalyze the cleavage and ligation of the two exons and removal of the intron (Moore, Query, and Sharp, 1993). This multistep process is depicted in Figure 1.2. There are several critical sequences involved in the splicing pathway. At the 5' end of the intron there is a splice donor sequence which binds the U1 snRNP. At the 3' end of the intron, the branch point sequence binds U2. There are two other important 3' sequences, the splice acceptor site and the polypyrimidine tract which are involved in the selection of the splice acceptor. UI and U2 then bind the U4/6 and U5 snRNPs which complex to form the spliceosome, the functional unit which removes introns from the primary transcript to produce a spliced RNA product (Adams, Rudner, and Rio, 1996; Moore, Query, and Sharp, 1993). This final product is then exported from the nucleus to the cytoplasm.

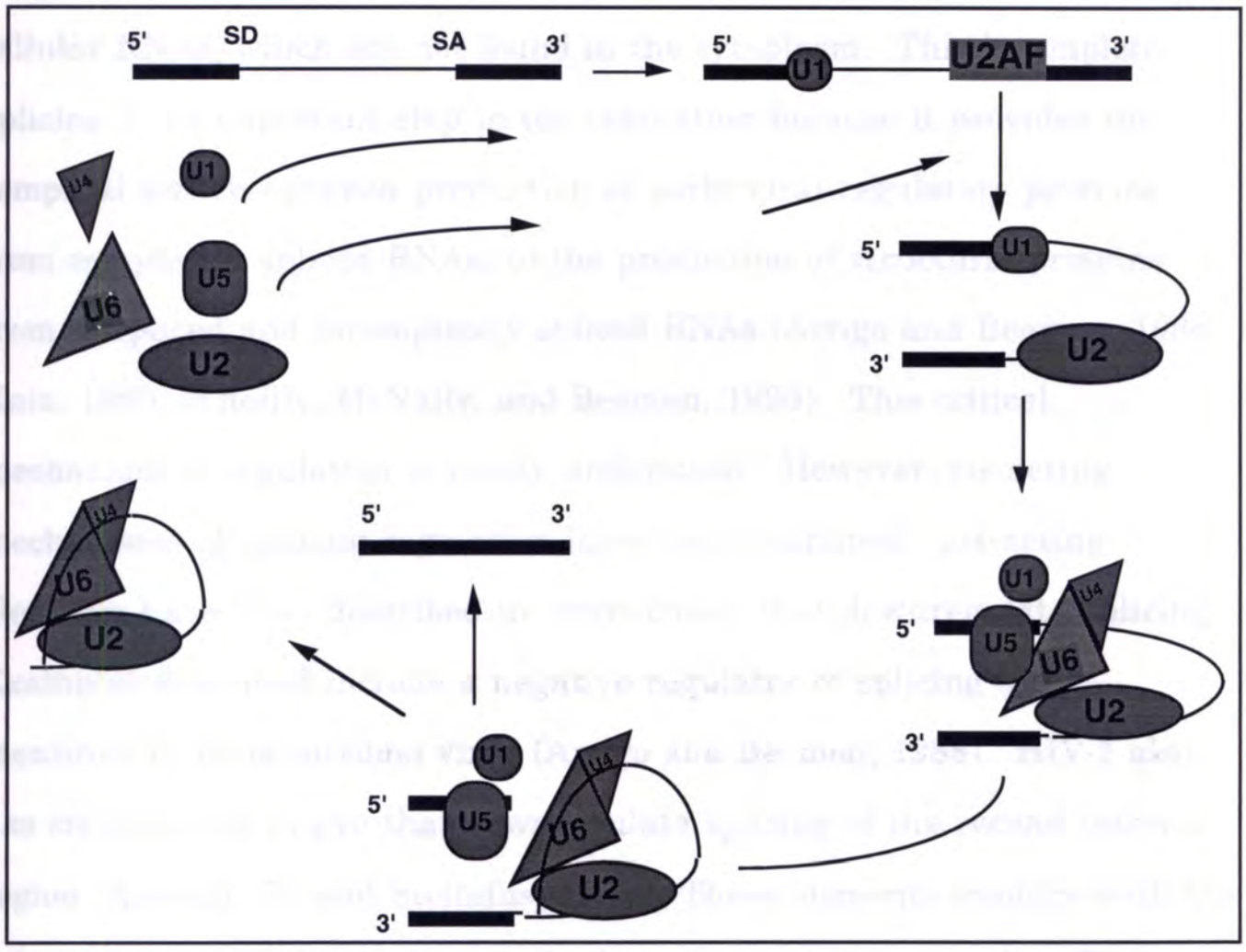


Figure 1.2 Splicing Pathway The process of splicing is shown with the various snRNPs binding to the primary RNA transcript followed by lariat formation, splice site cleavage and exon ligation the RNA to remove intronic sequences (Adapted from Moore by Justin Favaro) (Moore, Query, and Sharp, 1993).

The normal process of eukaryotic RNA splicing is inherently different from that of the splicing of retroviral transcripts. This is because of the fact that retroviral primary transcripts are multiply and variably spliced to produce different viral mRNAs. (see Figure 1.3 and 1.5). The primary transcript itself also encodes structural proteins as well as

functioning as the genomic RNA for progeny virions and is exported unspliced into the cytoplasm of the cell. This is in contrast to unspliced cellular RNAs, which are not found in the cytoplasm. This incomplete splicing is an important step in the regulation because it provides the temporal switch between production of early viral regulatory proteins from completely spliced RNAs, to the production of structural proteins from unspliced and incompletely spliced RNAs (Arrigo and Beemon, 1988; Katz, 1990; O'Reilly, McNally, and Beemon, 1995). This critical mechanism of regulation is poorly understood. However, *cis*-acting mechanisms of splicing regulation have been examined. *cis*-acting elements have been described in retroviruses that downregulate splicing. Examples described include a negative regulator of splicing (*nrs*) identified in Rous sarcoma virus (Arrigo and Beemon, 1988). HIV-1 also has cis elements in *env* that downregulate splicing of the second intronic region (Amendt, Si, and Stoltzfus, 1995). These elements combine with the suboptimal consensus sequence of splice sites and decrease the efficiency of a normally efficient process. The decrease in efficiency of splicing is critical to the ability to regulate this process.

After splicing, the cellular RNA is exported to the cytoplasm. Export of RNA is a regulated event that involves a series of steps that involve the binding of a number of cellular factors to facilitate transport across the nuclear membrane. The mechanisms of RNA export are currently a very active area of investigation as is the study of much of the regulation involving nuclear and cytoplasmic trafficking (Ullman, Powers, and Forbes, 1997). Elucidation of the mechanisms of transport and possible regulatory steps involved in the process will aid the understanding of the role of export in retroviral gene expression.

RNA metabolism is also regulated in the cytoplasm via interactions that regulate the stability and translational processing of the RNA. This is a brief description of the complex pathway of RNA processing. The importance of understanding the pathways of RNA processing can not be overemphasized. It is only through understanding the temporal processing of RNA that the possible regulation of those steps can be investigated. For now, a basic understanding of cellular RNA processing will allow the exploration of the gene regulation and the life cycle of a complex retrovirus.

The Human Immunodeficency Virus Type 1 (HIV-1) is a retrovirus which is the etiologic agent of the Acquired Immune Deficiency Syndrome (AIDS). (for review see Levy93) It is a member of the lentivirus genus of retroviruses. In a 10 kb genome, HIV-1 possesses at least 9 genes (Figure 1.3). These genes include *gag*, *pol*, and *env* common to all retroviruses. HIV-1 also has several regulatory genes- *tat* ,*rev* and *nef*. Other accessory regulatory genes include *vif*, *vpr*, *vpu* and *nef*. The *gag* gene encodes viral structural proteins. Translation of *gag* results in a polyprotein, which is subsequently cleaved by the viral protease into the component proteins- matrix, capsid and nucleocapsid. *Pol* encodes the viral enzymes (reverse transcriptase, integrase, RNase H and protease). *Env* encodes the envelope proteins of the virus (gp120 and gp41). gp120 is inserted into the cell membrane and thereby incorporated into virions during budding. In infection gp120 functions as a ligand for the CD4 receptor which along with a chemokine receptor, facilitates the binding of the virion to the cell. The fusion of the viral envelope to the cell during infection is facilitated by the action of gp41. *vif* is involved in the infectivity of the virus. *vpr* plays a role in the pathogenesis of the virus during infection. *vpu* plays a role in

virus assembly. Nef, a regulatory protein produced from a fully spliced RNA, has two major functions in the viral life cycle; including the downregulation of CD4 receptors and involvement in viral pathogenesis during infection.

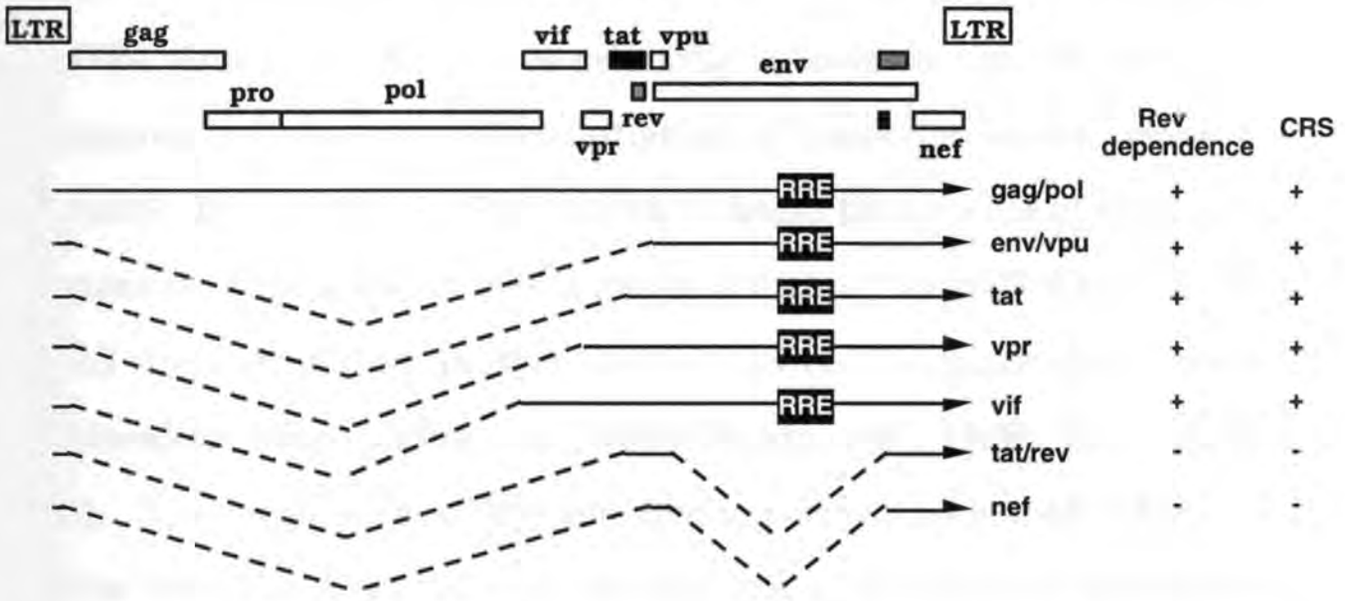


Figure 1.3 HIV-1 Genome and RNAs. The HIV-1 Genome and RNAs produced by the virus. The genome of HIV-1 is depicted along with the various RNAs produced through differential splicing of the full length RNA. Rev dependence and the presence of CRS are noted.

HIV-1 expresses at least 17 proteins from a 9.6kb genome. In order to do this; the virus utilizes the mechanisms outlined previously, which include the production of three different classes of RNA. These are produced through various combinations of the major splice donors and splice acceptors. These RNAs and their protein products are depicted at the bottom of Figure 1.3 along with the splicing events required for their formation.

The gene expression of HIV-1 is coordinated by the products of *tat* and *rev*. They share some common features. Both function in *trans* and are produced as early proteins produced from a fully spliced mRNA. Following translation both localize to the nucleus and bind target sites on RNA. After binding, *tat* and *rev* exert their effects on the RNA. Tat binds to the TAR element of the RNA and functions to increase transcription from the viral LTR. Tat serves primarily to increase the efficiency of elongation, but also may affect initiation of transcription (for review see Cullen95). It provides the "on" switch to begin the process of viral gene expression. Expression of HIV-1 genes is further controlled by Rev. Rev is a 19kd *trans* regulatory protein that is required for expression of viral structural proteins (Malim *et al.*, 1988; Malim *et al.*, 1989b; Sodroski *et al.*, 1986). The mechanism of Rev function is poorly understood. Many studies have found Rev to play a pivotal role at a number of points in viral gene expression. These points are the splicing, transport, stability and translation of mRNA. Rev binds to its recognition sequence, the Rev-response element (RRE), located in *env* (see Figure 1.3). Rev can be divided into two functional regions (Figure 1.4). The N-terminal region binds to the RRE as well as containing a nuclear localization signal. This region consists of a series of arginine residues (Berger *et al.*, 1991; Malim *et al.*, 1989a). Overlapping this domain is the multimerization region. Rev binds to the RNA as a monomer but subsequently multimerizes to achieve its effects (Olsen *et al.*, 1990). The C-terminal region of the protein contains an effector domain consisting of a series of leucines (Berger *et al.*, 1991; Malim *et al.*, 1989a). Mutations in this region abolish function of the protein but produce transdominant molecules. These transdominants are

capable of binding to the RRE and inhibit the function of wild type Rev (Arrigo, Heaphy, and Haines, 1992; Garrett and Cullen, 1992).

Figure 1.4 Structural Domains of Rev. The functional regions of the Rev protein are indicated in the above diagram. The multimerization domain overlaps the RRE binding domain near the N-terminal of the protein.

The RRE serves as a Rev binding site on the RNA. The RRE is a 234 nucleotide stem-loop structure normally located in *env* (Hadzopoulou *et al.*, 1989; Malim *et al.*, 1989b; Rosen *et al.*, 1988). It is present in unspliced and partially spliced RNAs but is removed from the fully spliced RNAs (see Figure 1.3). The location of the RRE, within the RNA, is not important; as intronic sites appear to function as well as exonic sites for Rev function (Campbell, Borg, and Arrigo, 1996; Malim *et al.*, 1989b). However, the effect of the RRE is orientation dependent (Hammarskjold *et al.*, 1989). The structure of the RRE is a complex folded RNA that has specific structural components. One of these is the high affinity binding site for Rev that has been ascribed to the IIb region of the element. Other regions and sequence are important in multimerization and further binding of Rev to the RRE (Kjems *et al.*, 1991; Mann *et al.*, 1994). Following binding of Rev to the region it can then exert its effects on the fate of the RNA.

Several theoretical mechanisms for Rev function have been proposed and tested. One is that Rev serves to enhance transport of the Rev-dependent RNAs, rescuing them from being trapped in the nucleus. This mechanism argues that Rev functions by binding to the RNA and then facilitating its transport from the nucleus to the cytoplasm. Experiments with Rev fused to BSA have shown functional export of the

proteins from the nuclei of Xenopus (Fischer *et al.*, 1994). One piece of data absent in this mechanism was the identification of the cellular cofactors required for Rev to function in this manner. This factor has now been identified by several groups as a nucleoporin-like protein. Rev associates with this nucleoporin-like protein, which theoretically enables transport of the Rev bound RNA. This factor was identified using the yeast two-hybrid system and contains similar motifs to nucleoporins, which are components of the nucleopore complex (Bogerd *et al.*, 1995; Fritz, Zapp, and Green, 1995; Stutz, Neville, and Rosbash, 1995). Consequently, this protein may function as a cellular cofactor for Rev to facilitate export of bound RNAs. However, additional work is required in order to establish a clear relationship between Rev, the nucleoporin-like protein and the mRNA.

Several lines of evidence suggest that Rev may directly affect splicing of HIV-1 RNAs in order to regulate production of the various species of HIV-1 RNA. An inhibition of splicing *in vitro* was demonstrated using a peptide containing the Rev RNA binding domain (Kjems, Frankel, and Sharp, 1991). This inhibition of splicing was not seen *in vivo*; therefore it is difficult to evaluate the relevance of this data. While certainly an attractive theory for Rev function, there is little evidence that Rev directly inhibits splicing to increase the production of unspliced and partially spliced RNA. It may be that the shift in production of RNAs is the result of changes in export, which indirectly change the levels of RNA or some combination of these effects.

Several studies have shown effects of Rev on translation of RNA. They have demonstrated a translational effect, in that in the absence of Rev, RNAs can accumulate in the cytoplasm, but are not translated.

Several groups have shown these effects in lymphoid cells (Arrigo and Chen, 1991) and fibroblast cells (D'Agostino *et al.*, 1992). The partially spliced messages accumulated in the cytoplasm, but were not translated. When examined by sucrose gradients, these untranslated RNAs were deficient in polysome formation. The fact that these RNAs could be translated *in vitro* strongly indicates a translational regulatory function for Rev. In the absence of Rev; Poly-A Binding Protein (PAB1) was not associated with the Rev-dependent messages. PAB1 binding has been shown to be critical for the association of the 60S ribosomal subunit and translation. At what stage Rev binding affects the association of PAB1 is unclear, but the failure to associate with PAB1 might be indicative of translational regulation of HIV-1 mRNAs by Rev.

A possible cofactor for Rev, eIF-5A, may shed light on a mechanism of translational regulation. Initial studies reported the interaction between Rev and eIF-5A. These studies were performed by crosslinking the proteins (Ruhl *et al.*, 1993). Additional studies using transdominant eIF-5A proteins have also suggested an interaction with Rev (Katahira *et al.*, 1995). This research is hampered by a lack of understanding of the function of eIF-5A. Without a clear understanding of the cellular role of eIF-5A, the importance of the data is difficult to estimate. While the level at which Rev functions and its various cellular cofactors are complex and poorly understood, the viral requirement for Rev remains absolute.

The answer to the question of why Rev is required for the expression of the Rev-dependent RNAs may lie in a series of ill-defined sequences known as Cis-acting Repressor Sequences (CRS). These sequences, in the absence of Rev, have a negative effect on the RNA and prevent its cytoplasmic accumulation. CRS are RNA sequences that are defined

functionally. When present in RNA, CRS negatively regulate the expression of the RNA. The second criterion required for definition as a CRS is that Rev must be able to rescue the RNA from this negative regulation through binding to the RRE. Two major groups of sequences have been implicated as CRS 1) sequences in *gag, pol* and *env* (intronic CRS) and 2) splice sites (see Figure1.3 and 1.5).

Several groups have worked to identify a series of CRS elements that are located in the *gag, pol* and *env* coding regions of the virus (Rosen *et al.*, 1988). These CRS have several factors in common. They are large (200+bp), function in the sense orientation, have no obvious sequence homology and require a large number of mutations to delete their negative function. The CRS located in *gag* is perhaps one of the best studied of this type. It is located in the p17 region and required 28 point mutations in the 270 bp element to inactivate. The inactivation of this CRS allowed for the discovery of another in the *gag-pol* region (Schwartz, Felber, and Pavlakis, 1992a). Other groups have also noted CRS in *pol* which appear similar to the CRS described in *gag*. The CRS is 260bp in size, located 3' in *pol* and was defined by deletion mutagenesis (Cochrane *et al.*, 1991). The nature of this element appears similar to that described for the *gag* CRS. The duplicative nature of these elements is certainly interesting. Why is there a need for so much redundancy in negatively regulating early expression of structural proteins? This question is largely unanswered.

Multiple elements in *env* have been described as CRS. These elements appear similar to those described in *gag* (Nasioulas *et al.*, 1994). After inactivation of an *env* CRS, other CRS were identified in *env*. This observation is similar to the identification of multiple elements in the *gag*-coding region. Several groups have reported that the RRE functions as a

CRS (Brighty and Rosenberg, 1994; Nasioulas *et al.*, 1994). This effect of the RRE is not seen in all systems, consequently the RRE may function as a CRS in only certain contexts. Studies of intronic CRS have not discovered a mechanism of function for the negative regulation of Rev dependent RNAs. Decreased stability, blocked export and translational inhibition have all been suggested as possible mechanisms for regulation of RNAs containing CRS.

The second major class of identified CRS elements is splice sites. These have been identified in several different systems as negatively regulating RNA expression. The addition of HIV-1 splice donors and acceptors to heterologous constructs induces Rev dependence (Chang and Sharp, 1989). The splice donors of HIV-1 are fairly good matches to the consensus sequence identified for splice sites (Mount, 1982). The splice acceptors are notably poor in HIV-1 and several groups have reported that suboptimal splicing events are necessary for Rev-responsiveness of the RNA (O'Reilly, McNally, and Beemon, 1995). Other groups have noted CRS effects of splice donors when placed in the 3' UTR of papillomavirus constructs (Barksdale and Baker, 1995) and of the 5' splice donor of Equine Infectious Anemia Virus (Tan *et al.*, 1996). The major 3' splice donor has been the subject of previous investigation and mediates the Rev dependence of *env* constructs. When the splice donor is mutated and an excisable intron was placed upstream, *env* expression became Rev independent (Hammarskjold *et al.*, 1994). These studies indicate that splice donors can function as CRS and that the splicing event must be suboptimal in order to be regulated by Rev. The mechanisms of CRS function and Rev rescue of Rev dependent RNA are still largely unknown. Some groups have suggested instability, impaired nuclear export or

translational effects as the means by which RNAs are negatively regulated in the absence of Rev, but these data are far from conclusive (Dayton, 1996).

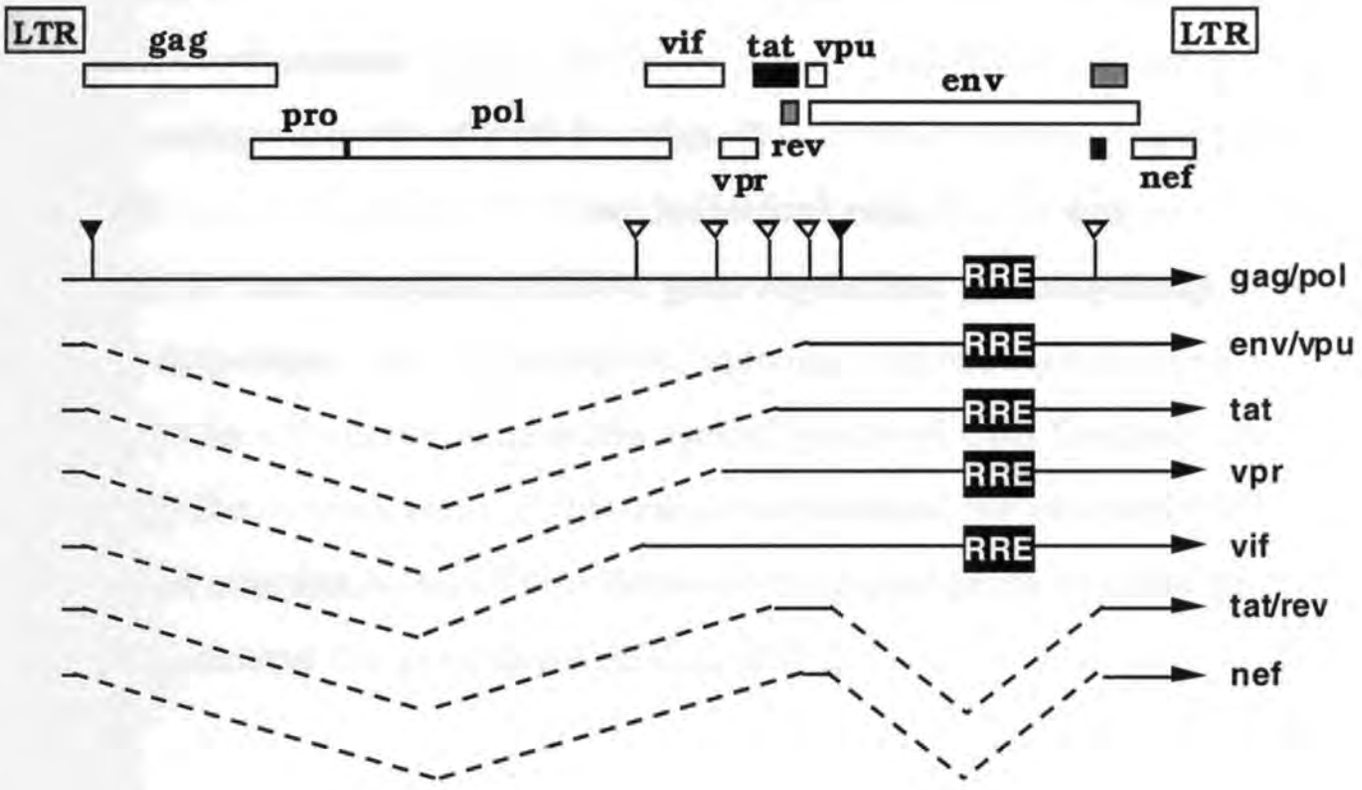


Figure 1.5 Splice Sites of HIV-1
The major splice sites of HIV-1 are depicted. Splice donors are shown as black triangles and splice acceptors as white triangles. These sites are used in combination to produce the viral RNAs.

CRS facilitate the regulation of HIV-1 mRNA by Rev. However, the molecular mechanisms of this regulation remain undefined. In order to begin to explore the role of CRS function, it is critical to dissect apart the Rev-RRE axis of expression so those individual components can be examined for their function in HIV-1 gene regulation. Consequently, it is critically important that transcription, splicing, export, stability and translation be considered as possible control points of CRS function. In examining the various steps of this axis of expression, we can begin to tease apart complex relationships between the components in order to better understand the gene expression of HIV-1.

Chapter 2

Involvement of Human Immunodeficiency Type-1 Splice Sites in the Cytoplasmic Accumulation of Viral RNA

Introduction

Splice sites are important elements involved in the Rev-responsiveness of HIV-1 mRNAs. This is true for several reasons. One is the variation in patterns of RNA produced in the presence and absence of Rev. Rev mediates the shift from the production of fully spliced RNA to the production of predominantly unspliced and partially spliced RNAs. The fact that splice donors can function as CRS also implicates splice site involvement in the Rev responsiveness of RNAs. By negatively regulating the RNA, splice donors serve as sequences (or at least some of the sequences) that establish the viral requirement for Rev. In order to explore this activity, we first wish to isolate splice sites to examine their role as regulatory elements involved in the cytoplasmic accumulation of HIV-1 mRNAs. A series of experiments with isolated splice sites will serve to characterize the function of these elements in HIV-1 gene expression. The constructs used to examine splice site function will contain deletions for known intronic CRS so that isolated splice donors and splice acceptors can be examined to determine their role in Rev responsiveness of HIV-1 mRNA.

The expression of HIV-1 structural proteins is regulated at the level of splicing, transport, stability, and translation by the Rev protein (Arrigo and Chen, 1991; Cochrane *et al.*, 1991; D'Agostino *et al.*, 1992; Feinberg *et al.*, 1986; Felber *et al.*, 1989; Hadzopoulou *et al.*, 1989; Hammarskjold *et al.*, 1989; Lawrence *et al.*, 1991; Malim *et al.*, 1988; Schwartz, Felber, and Pavlakis, 1992a; Sodroski *et al.*, 1986). The Rev protein binds to an RNA binding site termed the Rev-responsive element (RRE) present within the RNAs that it regulates (Hadzopoulou *et al.*, 1989; Malim *et al.*, 1989b; Rosen *et al.*, 1988). This regulation may involve alteration of the cellular

pathway of the RNA (Fischer *et al.*, 1995). The Rev-dependent RNAs require Rev for protein expression and therefore must contain negative sequences, the effect of which is overcome by the Rev protein (for review see Dayton 1996). The Rev-independent *tat*, *rev*, and *nef* genes presumably do not contain these sequences. These negative sequences have been termed cis-acting repressor sequences, CRS, and have been characterized in a variety of experimental systems. Splice sites have been implicated in Rev regulation by several groups. Introduction of sequences containing the HIV-1 *tat* 3' splice donor or splice acceptor sequences (i.e., those splice sites that join the Tat coding exons) into a construct containing a ß-globin intron conferred Rev-responsiveness to the construct (Chang and Sharp, 1989). Rev-responsiveness could also be conferred by mutation of the ß-globin splice site consensus sequence within this construct. The CRS potential of splice donor sequences has been demonstrated by the Rev-responsiveness of papillomavirus constructs containing a splice donor in the 3' untranslated region (Barksdale and Baker, 1995). The major 5' splice donor of equine infectious anemia virus functions as a CRS within EIAV Gag-producing constructs, although it functions in a location-dependent manner (Tan *et al.*, 1996). Within HIV-1 Env-expressing constructs, splice sites and U1 interaction have been shown to be important for the regulation of Env expression by Rev (Lu *et al.*, 1990). Furthermore, the expression of Env could be made Rev-independent by placing a complete intron upstream of *env* and removing known splice donor sequences (Hammarskjold *et al.*, 1994).

Other sequences within HIV-1, distinct from splice sites, have also been shown to exhibit CRS activity. The fusion of various fragments of the *gag*, *pol*, and *env* genes to a CAT reporter construct resulted in Rev-

responsiveness (Cochrane *et al.*, 1991; Rosen *et al.*, 1988). The CRS within *pol* elicited an orientation-dependent 30-fold negative effect on CAT activity that could be rescued by Rev and the RRE. This CRS is located within HIV-1 sequences that are not known to contain any splice sites. The effect appeared to be at the mRNA utilization level in that the subcellular distribution of RNA was not substantially affected. Fusion of regions of both *gag* and *pol* to reporter constructs also demonstrated at least two orientation-dependent CRS that functioned posttranscriptionally (Maldarelli, Martin, and Strebel, 1991). These sequences do not contain known splice sites and were not substantially affected by the addition of splice sites to the constructs. Fusions of *gag* to a *tat*-expression vector were used to analyze CRS activity within *gag* (Schwartz, Felber, and Pavlakis, 1992a). A less than 300 base pair fragment of *gag* was capable of exerting an effect on the total RNA production, which was attributed to differences in RNA stability. Mutational inactivation of this CRS required 28 point mutations, demonstrating the complex nature of this CRS (Schwartz *et al.*, 1992). In the absence of known splice sites, *env* expression was shown to be Rev-dependent in mammalian cells using Env expression vectors (Nasioulas *et al.*, 1994). Localization of CRS within *env* identified many regions capable of exerting modest negative activity, which together exerted activity that is more dramatic. Deletion of the RRE led to Rev-independent expression of Env from subgenomic *env* constructs, indicating that a CRS may overlap RRE sequences (Brighty and Rosenberg, 1994; Churchill *et al.*, 1996). The complexity of the interaction between the *gag* CRS and splicing was recently demonstrated (Mikaelian *et al.*, 1996). The *gag* CRS in conjunction with the RRE was capable of conferring Rev-responsiveness

on a ß-globin construct. Mutations in splice sites rendered these constructs Rev-independent. Therefore, it appears that CRS may function at a variety of levels to decrease the expression of RNAs containing these elements.

In this chapter, we investigated the involvement of HIV-1 splice sites in Rev regulation through sequential mutagenesis of an infectious proviral clone. In order to avoid the complexities of analyzing multiple types of CRS, the majority of intronic sequences were deleted, retaining the splice sites necessary for the production of the multiply spliced RNAs. Analysis of the ratio of cytoplasmic to nuclear HIV-1 RNAs, in the presence or absence of Rev or the RRE, using these constructs served as an indicator of CRS activity.

Materials and Methods

Plasmid Constructions

The BI construct was constructed from pYKJRCSF/EBV- (Arrigo and Chen, 1991) using DR mutagenesis (Gatlin *et al.*, 1995). *gag*, *pol*, *vif* and part of *vpr* were deleted (nt790-5710), leaving the major splice donor and Gag methionine intact. *env* and *vpu* were removed by deleting from just downstream of the 3' splice donor for *tat*, *rev*, and *nef* to just upstream of the most 3' splice acceptor (nt6145-8307). The BI construct retains the splice sites for the generation of *tat*, and is deleted for the majority of *tat* intronic sequences. The BIR construct was generated by insertion of an RRE-containing *Sal*I fragment into the *Xho*I site within the *nef* sequences of the BI construct (Campbell, Borg, and Arrigo, 1996). FI and FIR constructs were generated using DR mutagenesis. The second *tat/rev* intron was deleted in frame from the BI and BIR constructs using the following oligonucleotide primers: 5'-TGCTTTGGTAGAGAAACTTGATGAGTCTGAC-3' and 5'-ACCCTCCTCCCAGCAACGAG-3'. This deletion creates an in-frame fusion of the two *tat* coding exons. SI and SIR constructs were generated using DR mutagenesis. The first *tat* intron was deleted from BI and BIR constructs using the following oligonucleotide primers: 5'-CAGTCGCCGCCCCTCGCCTC-3' and 5'-AATTGGGTGTCGACATAGCAGAATAGGC-3'.

The SD2R construct was generated by restriction digestion of the SIR construct with *Bam*HI and religation. The SD1 construct was generated using PCR and contains sequences from nt153-793, and nt9036-9736 from pYKJRCSF, including the sequences encompassing the major 5' HIV-1

splice donor. These sequences were cloned into pGem2 (Promega). The SD1R construct was generated by insertion of the RRE-containing *Spe*I fragment (Campbell, Borg, and Arrigo, 1996) into the SD1 construct. The NI construct was generated by the cloning of a PCR fragment into the SD1 construct. A 160 bp fragment was amplified using the BI construct as a template and the oligonucleotide primers: 5'-ACAACCATGGAGGCTAGCTAGGGAACCCACTG-3' and 5'-AACAGAATTCTCGCTTTCCGGTCCCTGTTC-3'. PCR conditions were as previously described (Arrigo, Heaphy, and Haines, 1992). The PCR product and the SD1 construct were digested with *Nco*I and *Eco*RI. The PCR product was inserted into the SD1 construct to generate NI using standard molecular techniques. The NI construct contains a deletion of almost all sequences between the splice donor and the polyadenylation signal, retaining 78 base pairs between these two elements. ΔNI was generated from the NI construct using DR mutagenesis and the following oligonucleotide primers: ΔSD1, 5'-AGAACCGCGGACGCCGAAATTTTGACTAGCGGAGGC-3' and ΔSD1A, 5'-ACTACCGCGGGTCGCCGCCCCTCGCCTCTTGC-3'. This results in the deletion of the consensus splice donor site, TGGTGAGT, from the NI construct and the insertion of the sequence CCGCGG.

Cell Culture and Transfection

COS cells were maintained in Iscove's medium supplemented with 10% fetal calf serum. Transfections were performed as described previously (Arrigo *et al.*, 1989). A total of 50 μg of plasmid DNA was electroporated for each sample. Cotransfections were performed with the indicated construct without or without the addition of pSVtat and/or

pCMV Rev (Lewis *et al.*, 1990). A filler construct (pcDNA3, Invitrogen) was used to normalize the total amount of DNA to 50μg. A confluent T175cm^2 flask of cells was used for approximately six transfections. RNA was harvested approximately 40 hours post-electroporation.

RNA fractionation and analysis

Cells were fractionated as previously described (Favaro and Arrigo, 1997). Isolation of RNA and RT-PCR was performed as previously described (Arrigo and Chen, 1991; Arrigo *et al.*, 1989; Arrigo *et al.*, 1990). RT-PCR was performed for 20 cycles of amplification for all RNAs except that of the total RNA produced by NI and ΔNI which was amplified for 15 cycles. Most of the oligonucleotide primers used for detection have been previously described (Arrigo and Chen, 1991; Arrigo *et al.*, 1989; Arrigo *et al.*, 1990). The unspliced RNAs produced by the BI/BIR, FI/FIR, SD1/SD1R, and NI/ΔNI constructs was detected with the LA8/Gag-Nco oligonucleotide primers. The unspliced RNAs produced by the SI/SIR and SD2/SD2R constructs were detected with the LA45/AA821 oligonucleotide primers. The spliced RNAs produced by the BI/BIR and SI/SIR constructs were detected with the LA45/LA41 oligonucleotide primers. The spliced RNAs produced by the FI/FIR constructs were detected with the LA8/LA23 oligonucleotide primers. Oligonucleotide primers LA23 (5'-GCCTATTCTGCTATGTCGACACCC-3'; nt5815-5792) and Gag-Nco (5'-CGCACCCATGGCTCTCCTTCTAG-3'; nt797-775) are antisense primers homologous to sequences downstream of the most 5' *tat* splice acceptor and major 5' splice donor, respectively. Total RNA produced by the SD1R, SD2R, and NI/ΔNI constructs was detected with an oligonucleotide pair located upstream of the major 5' splice donor: LDR1 (5'-

GAACAGGGACCGGAAAGCG-3'; nt644-662) and α-LDR1 (5'-GCGCGCTTCAGCAAGCCGAG-3'; nt716-697).

Quantitation of the ratio of cytoplasmic to nuclear RNA was performed by scanning autoradiographs with a Lacie DTP Silverscan III scanner. Individual bands were quantitated with NIH Image 1.59. Analysis was performed on a Macintosh 7100/80 computer using the public domain NIH Image program (developed at the U.S. National Institutes of Health and available from the Internet by anonymous FTP from zippy.nimh.nih.gov or on floppy disk from the National Technical Information Service, Springfield, Virginia, part number PB95-500195GEI). The cytoplasmic fractions (Cyto 1 and Cyto 2) were added together and compared to the nuclear fraction to determine the ratio of cytoplasmic to nuclear RNA. The level of nuclear RNA was arbitrarily set to a level of 1.

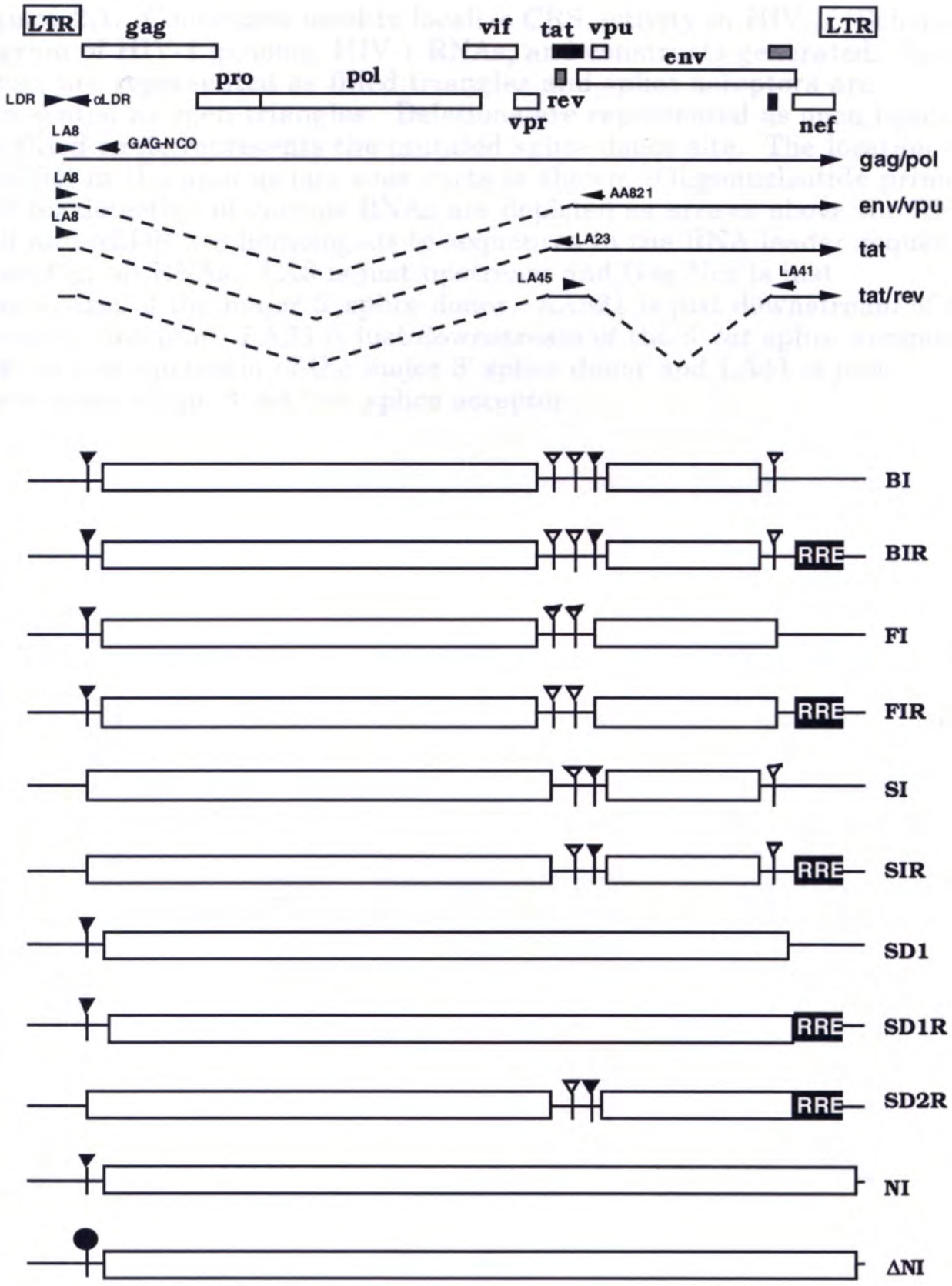
Results

Deletion of the majority of intronic sequences from HIV-1 does not eliminate CRS activity.

CRS activity has been attributed to intronic sequences within *gag*, *pol* and *env*. Starting with an infectious clone of HIV-1, the majority of intronic sequences were deleted. These deletions remove most of the sequences known to contain CRS activity. The resultant construct, BI, retains minimal amounts of both HIV-1 *tat* introns (Fig. 2.1). The BI construct retains 145 nucleotides of 5' and 150 nucleotides of 3' intronic sequences. The RNA produced by the BI construct should retain the

ability to splice both of these minimal introns and should produce all three classes of HIV-1 RNA (unspliced, partially spliced and fully spliced). In addition, this construct should be capable of producing the Tat and Rev regulatory proteins. This construct did not contain the RRE, which has been shown to exhibit CRS activity in some experimental systems (Brighty and Rosenberg, 1994; Churchill *et al.*, 1996). The RRE was inserted into the *nef* gene of this construct to generate the construct BIR (Fig. 1). The presence of the RRE at this site has been previously shown to function in a similar manner to the normal location of the RRE (Campbell, Borg, and Arrigo, 1996). Since the BIR construct both produces Rev and contains the RRE, comparison of the distributions of RNAs produced by this construct with the distributions of RNAs produced by the BI construct should provide a good indication of the presence of CRS activity in the BI construct. The BI and BIR constructs were analyzed by transient transfection into COS cells. The cells were separated into three subcellular fractions (Favaro and Arrigo, 1997). RNA was isolated and analyzed using a quantitative RT-PCR procedure (Arrigo *et al.*, 1989). The results are shown in figure 2.2. The autoradiographs were quantitated to determine the ratio of cytoplasmic to nuclear RNA (C:N). A ratio of 1:1 indicates that 50% of the RNA was found in the cytoplasmic fractions and 50% was found in the nuclear fraction. In the absence of the RRE, the unspliced RNA produced by the BI construct had a predominantly nuclear distribution, as evidenced by a C:N ratio of 0.2:1. In the presence of the RRE, the unspliced RNA produced by the BIR construct had a predominantly cytoplasmic distribution, as evidenced by a C:N ratio of 4.5:1. The fully spliced RNA accumulated predominantly in the cytoplasm, regardless of the presence or absence of the RRE. The

ratios of cytoplasmic to nuclear spliced RNA produced by the BI and BIR constructs were 20:1 and 16:1, respectively. The difference in the C:N ratio of unspliced RNA produced by the BI and BIR constructs indicated that the unspliced RNA produced by the BI construct retained substantial levels of CRS activity.



Key

Deletion

Splice Donor

Splice Acceptor

Mutated Splice Donor

Figure 2.1. Constructs used to localize CRS activity in HIV-1. Schematic diagram of HIV-1 genome, HIV-1 RNAs, and constructs generated. Splice donors are represented as filled triangles and splice acceptors are represented as open triangles. Deletions are represented as open boxes. The filled circle represents the mutated splice donor site. The location of the RRE in the appropriate constructs is shown. Oligonucleotide primers used for detection of various RNAs are depicted as arrows above the RNAs. LDR and αLDR are homologous to sequences in the RNA leader sequence present in all RNAs. LA8 is just upstream and Gag-Nco is just downstream of the major 5' splice donor. AA821 is just downstream of the *env* splice acceptor. LA23 is just downstream of the 5' *tat* splice acceptor. LA45 is just upstream of the major 3' splice donor and LA41 is just downstream of the 3' *tat*/*rev* splice acceptor.

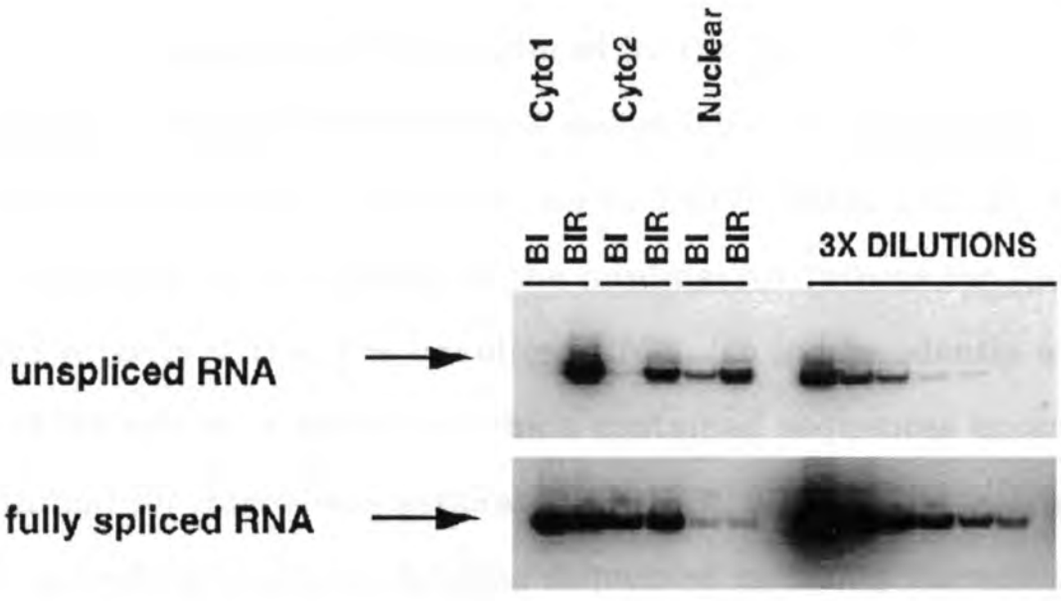


Figure 2.2. Effect of deletion of the majority of intronic sequences on CRS activity. Results of RT-PCR analysis, using fractionated RNAs (Cyto 1, Cyto 2 and Nuclear) from COS cells transfected with 50μg of the indicated constructs. Cyto1 and Cyto 2 fractions were generated using a buffer containing 0.05% NP40 and 0.65% NP40, respectively. The remaining pellet was designated the Nuclear fraction. Detection of unspliced and fully spliced RNAs is shown. Threefold dilutions of RNA from transfected cells were assayed to demonstrate the quantitative nature of this procedure.

Presence of a minimal intron maintains CRS activity

The unspliced RNA produced by the BI construct contained CRS activity, although this construct was deleted for the majority of HIV-1 intronic sequences. Therefore, we wished to independently examine constructs containing each of the minimal *tat* introns for the retention of CRS activity within the unspliced RNA. To independently examine the first *tat* intron, a construct which contained sequences encoding the first minimal *tat* intron was generated (FI; Fig. 2.1). This construct was generated by precisely deleting sequences encoding the second *tat* intron from the BI construct, conserving *tat* and *rev* reading frames. The FI construct should produce an unspliced RNA and a spliced RNA encoding Tat and Rev proteins. The FIR construct was generated through insertion of an RRE into the *nef* gene of the FI construct (Fig. 2.1). Both FI and FIR constructs were transfected into COS cells and assessed by RT-PCR analysis of fractionated RNAs (Fig. 2.3). In the absence of the RRE, the unspliced RNA produced by the FI construct had a predominantly nuclear distribution, as evidenced by a C:N ratio of 0.4:1. In the presence of the RRE, the unspliced RNA produced by the FIR construct had a predominantly cytoplasmic distribution, as evidenced by a C:N ratio of 2.0:1. The fully spliced RNA accumulated predominantly in the cytoplasm, regardless of the presence or absence of the RRE. The ratios of cytoplasmic to nuclear spliced RNA produced by the FI and FIR constructs were 4.4:1 and 3.8:1, respectively. The difference in the C:N ratio of unspliced RNA produced by the FI and FIR constructs indicated that the

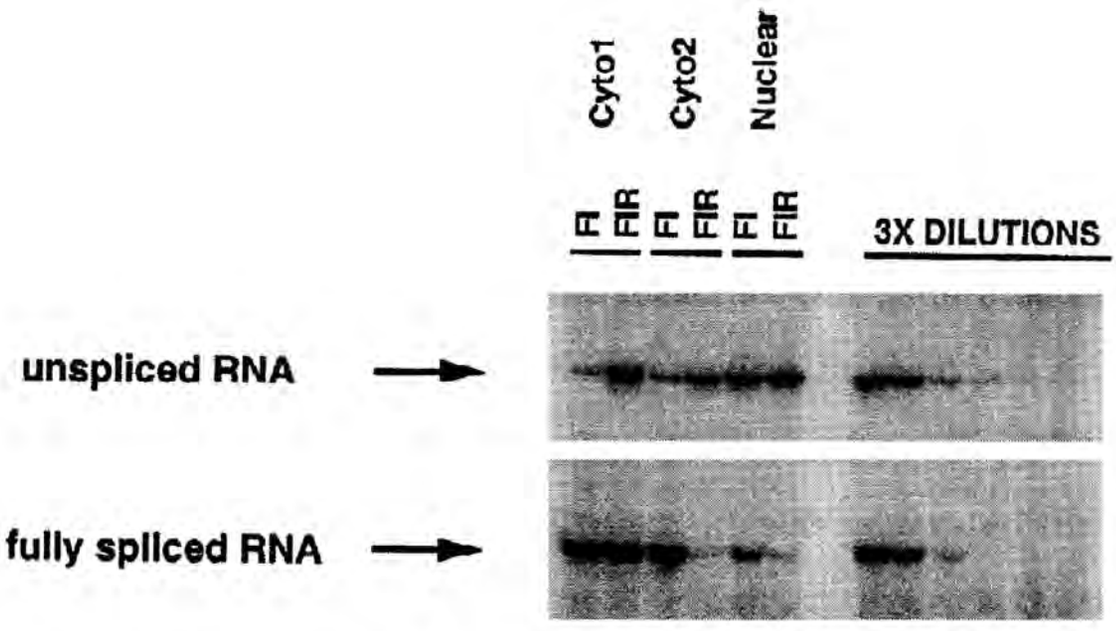


Figure 2.3. Effect of the first minimal *tat* intron on CRS activity. Results of RT-PCR analysis, using fractionated RNAs (Cyto 1, Cyto 2 and Nuclear, as in Fig. 2) from COS cells transfected with 50μg of the indicated constructs. Detection of unspliced and fully spliced RNAs is shown. Threefold dilutions of RNA from transfected cells were assayed to demonstrate the quantitative nature of this procedure.

unspliced RNA produced by the FI construct retained substantial levels of CRS activity.

To independently examine the second *tat* intron for CRS activity, a construct containing sequences encoding the second minimal *tat* intron was generated (SI; Fig. 2.1). This construct was generated by deletion of sequences encoding the first *tat* intron. The SI construct should produce an unspliced RNA encoding the Tat protein and a spliced RNA encoding Tat and Rev proteins. The SIR construct was generated through insertion of an RRE into the *nef* gene of the SI construct (figure 2.1). Both the SI and SIR constructs were transfected into COS cells and assessed for CRS activity by RT-PCR analysis of fractionated RNAs. The results are shown in figure 2.4. In the absence of the RRE, the unspliced RNA produced by the SI construct had a predominantly nuclear distribution, as evidenced by a C:N ratio of 0.8:1. In the presence of the RRE, the unspliced RNA produced by the SIR construct had a predominantly cytoplasmic distribution, as evidenced by a C:N ratio of 2.4:1. The fully spliced RNA accumulated predominantly in the cytoplasm, regardless of the presence or absence of the RRE. The ratios of cytoplasmic to nuclear spliced RNA produced by the SI and SIR constructs were 20:1 and 16:1, respectively. The difference in the C:N ratio of unspliced RNA produced by the SI and SIR constructs indicated that the unspliced RNA produced by the SI construct retained substantial levels of CRS activity.

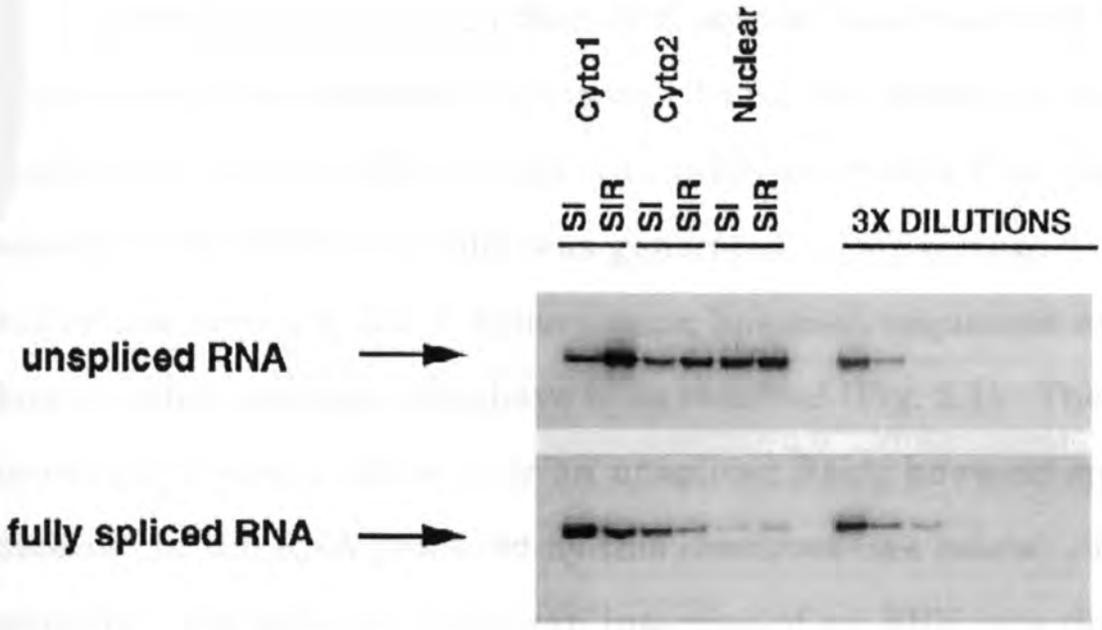


Figure 2.4. Effect of the second minimal *tat* intron on CRS activity. Results of RT-PCR analysis, using fractionated RNAs (Cyto 1, Cyto 2 and Nuclear, as in Fig. 2) from COS cells transfected with 50μg of the indicated constructs. Detection of unspliced and fully spliced RNAs is shown. Threefold dilutions of RNA from transfected cells were assayed to demonstrate the quantitative nature of this procedure.

Deletion of utilized splice acceptors does not eliminate CRS activity

Results had indicated that CRS activity was exhibited by RNAs containing either minimal *tat* intron. To further define the sequences responsible for the CRS activity within RNAs produced by the FI construct, the SD1 construct was generated. This construct retains the sequences encoding the 5' splice donor; however, sequences encoding all known splice acceptor sites have been removed (Fig. 2.1). The SD1 construct should produce only an unspliced RNA; however splicing was detected in the RNA produced by this construct (see below). The SD1R construct was generated through insertion of an RRE into the SD1 construct (Fig. 2.1). SD1 and SD1R constructs were transfected into COS cells and assessed by RT-PCR analysis of fractionated RNAs using oligonucleotide primers that flanked the splice donor sequences to detect unspliced RNA. As these constructs do not produce Tat, all transfections included a Tat-producing plasmid. These constructs also do not produce Rev; therefore, transfections were performed in the presence or absence of a Rev-producing plasmid. The results are shown in figure 2.5, upper panel. The unspliced RNA produced by the SD1 construct, in the absence of Rev, was found predominantly in the nuclear fraction as evidenced by a C:N ratio of 0.6:1. The C:N ratio of unspliced RNA produced by the SD1 construct was not affected by the addition of Rev, as evidenced by a C:N ratio of 0.4:1. In the absence of Rev, the SD1R construct exhibited a similar RNA profile to that of the SD1 construct, as evidenced by a C:N ratio of 0.6:1. However, in the presence of Rev, the ratio of cytoplasmic to nuclear unspliced RNA produced by the SD1R construct was increased to a ratio of 5.5:1. These results indicated that the SD1R construct,

containing the major HIV-1 5' splice donor, retained CRS activity within the unspliced RNA.

Although no known splice acceptor sequences were present within the SD1R construct, it was still possible that a cryptic spice acceptor might have been utilized in conjunction with the 5' splice donor. To determine whether the splice donor was utilized in the RNA produced by the SD1R construct, we further examined the RNA produced by this construct using oligonucleotide primers specific for sequences upstream of the splice donor to detect total RNA, both spliced and unspliced (Fig. 2.5, lower panel). In contrast to results using primers that detected only unspliced RNA, the use of this set of primers indicated that Rev had little effect on the ratio of cytoplasmic to nuclear RNA produced by the SD1R construct. In fact, the bulk of the total RNA was found in the cytoplasmic fractions regardless of the presence or absence of Rev as evidenced by C:N ratios of 3.6:1 in the presence of Rev and 6.6:1 in the absence of Rev. This difference in C:N ratio was not seen in other experiments. These results indicated that the splice donor site within the RNA produced by the SD1R construct was utilized in combination with a cryptic splice acceptor site(s).

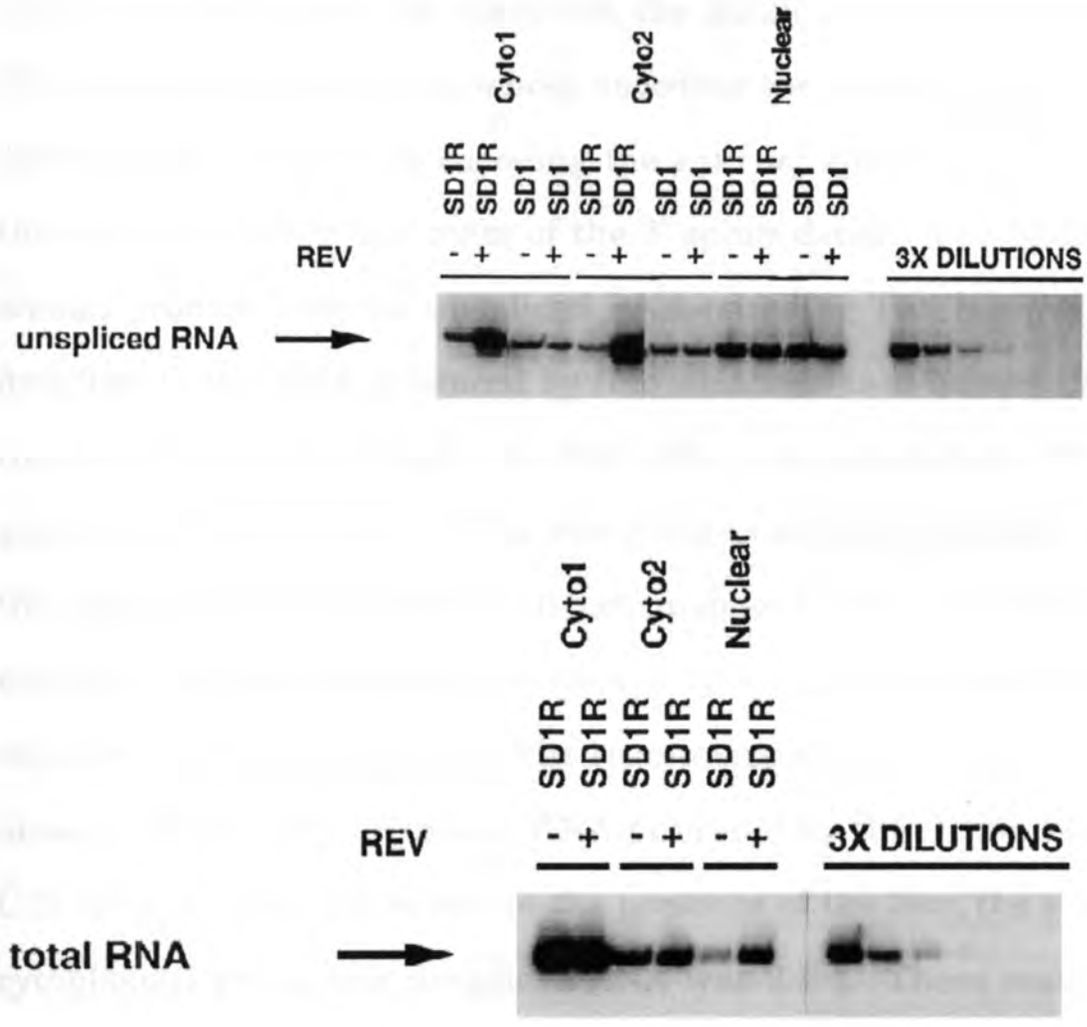


Figure 2.5. Effect of deletion of known splice acceptors and retention of the major 5' splice donor on CRS activity. Results of RT-PCR analysis, using fractionated RNAs (Cyto 1, Cyto 2 and Nuclear, as in Fig. 2) from COS cells transfected with 30μg of the indicated constructs. Detection of unspliced (upper panel) and total (lower panel) RNAs is shown. Threefold dilutions of RNA from transfected cells were assayed to demonstrate the quantitative nature of this procedure. All transfections also contained 10 μg of a Tat-producing construct. The presence or absence of 10μg of a Rev-producing construct is indicated(+ or-).

To better elucidate the CRS activity present within the unspliced RNA produced by the SI construct, the SD2R construct was generated. This construct retains sequences encoding the major 3' HIV-1 splice donor. Although the sequences encoding the *env/nef* splice acceptor are retained, they are positioned upstream of the 3' splice donor. The SD2R construct should produce only an unspliced RNA encoding Tat; however splicing was detected in the RNA produced by this construct (see below). The SD2R construct was transfected into COS cells and assessed by RT-PCR analysis of fractionated RNAs using oligonucleotide primers that flanked the splice donor sequences to detect unspliced RNA. As this construct does not produce Rev, transfections were performed in the presence or absence of a Rev-producing plasmid. The results are shown in figure 2.6. In the absence of Rev, the unspliced RNA produced by the SD2R construct had a C:N ratio of 0.5:1. However, in the presence of the Rev, the ratio of cytoplasmic to nuclear unspliced RNA was 2.8:1. These results indicated that the unspliced RNA produced by the SD2R construct retained CRS activity.

The RNA produced by the SD2R construct was further analyzed using oligonucleotides directed against sequences upstream of the splice donor to detect total RNA produced by the construct (Fig. 2.6). In contrast to results using primers that detected only unspliced RNA, the use of these primers indicated that Rev had little effect on the distribution of RNA produced by the SD2R construct as evidenced by C:N ratios of 3.5:1 in the absence of Rev and 4.4:1 in the absence of Rev. These results indicated that the splice donor site within the SD2R construct was utilized in combination with a cryptic splice acceptor site(s).

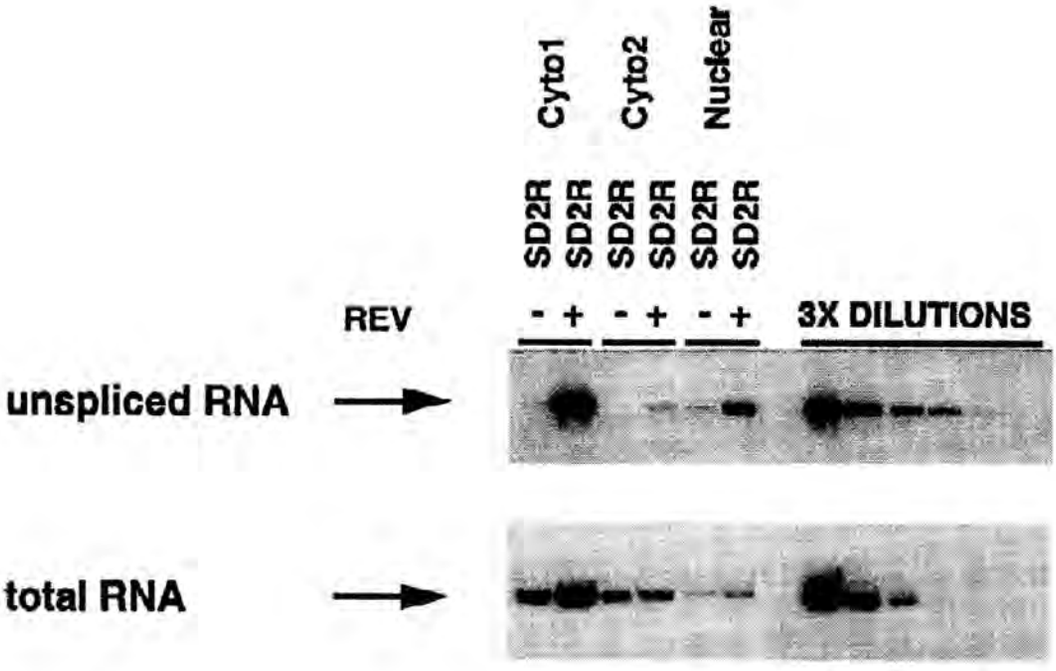


Figure 2.6. Effect of deletion of known downstream splice acceptors and retention of the major 3' splice donor on CRS activity. Results of RT-PCR analysis, using fractionated RNAs (Cyto 1, Cyto 2 and Nuclear, as in Fig. 2) from COS cells transfected with 25μg of the SD2R construct. Detection of unspliced and total RNAs is shown. Threefold dilutions of RNA from transfected cells were assayed to demonstrate the quantitative nature of this procedure. The presence or absence of 25μg of a Rev-producing construct is indicated (+ or -).

An isolated splice donor functions as a CRS.

Results obtained using the SD1R and SD2R constructs indicated that, in both cases, the splice donors were being utilized in conjunction with cryptic splice acceptors. The NI construct was generated through deletion of the bulk of sequences between the splice donor and the polyadenylation signal, leaving only 78 base pairs between these elements. Due to the lack of sequences downstream of the splice donor, we anticipated that the RNA produced by the NI construct would be extremely limited in its utilization of the major 5' splice donor. To elucidate the role of the major 5' splice donor in the manifestation of CRS activity, site-directed mutagenesis of this site within the NI construct was performed. The consensus splice donor sequence within the NI construct was mutated to generate the NI construct (Fig.2.1). NI and ΔNI constructs were transfected into COS cells in duplicate. Since these constructs do not produce Tat, all transfections included a Tat-producing plasmid. Fractionated RNAs were assessed by RT-PCR analysis using oligonucleotide primers that flanked the splice donor sequences. The results are shown in figure 2.7. The unspliced RNA produced by the NI construct was detected predominantly in the nuclear fraction as evidenced by a C:N ratio of 0.2:1-0.3:1. Mutation of the splice donor site (ΔNI) resulted in an increase in the ratio of cytoplasmic to nuclear unspliced RNA as evidenced by a C:N ratio of 1.0:1-1.3:1. These results indicated that the 5' splice donor was functioning as a CRS within the unspliced RNA produced by the NI construct. To determine the extent to which the 5' splice donor was utilized, the RNA produced by the NI and ΔNI

constructs was further analyzed using oligonucleotides directed against sequences upstream of the splice donor to detect total RNA, both spliced and unspliced (Fig. 2.7). The ratio of cytoplasmic to nuclear total RNA produced by the NI construct was altered in that the C:N ratio was increased to 1.0:1-1.5:1 . These results indicated that splicing was not completely eliminated within the RNA produced by the NI construct.

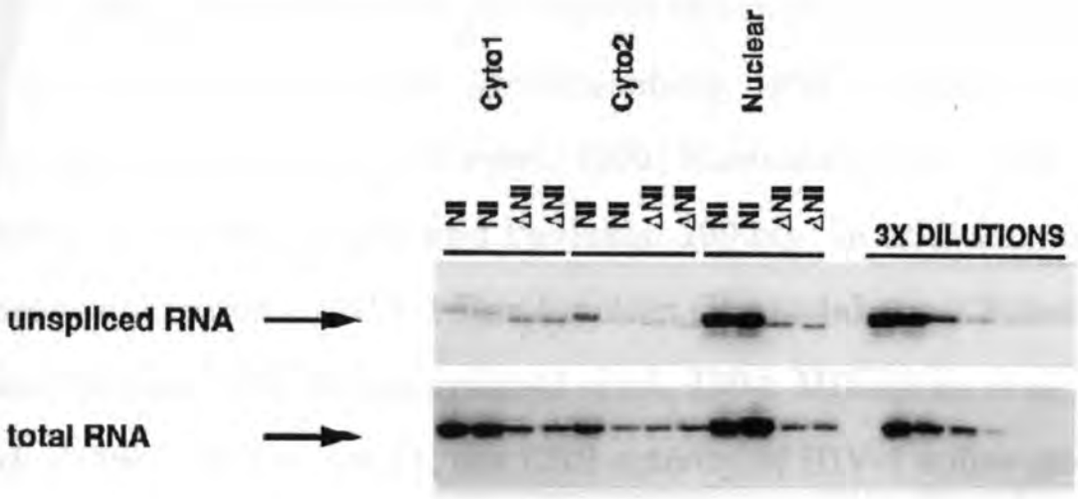


Figure 2.7. The major 5' HIV-1 splice donor functions as a CRS. Results of RT-PCR analysis, using fractionated RNAs (Cyto 1, Cyto 2 and Nuclear, as in Fig. 2) from COS cells transfected with 40μg of the indicated constructs in duplicate. Detection of unspliced and total RNAs is shown. Threefold dilutions of RNA from transfected cells were assayed to demonstrate the quantitative nature of this procedure. All transfections also contained 10 μg of a Tat-producing construct.

Discussion

CRS have been identified throughout the intronic sequences of HIV-1, in *gag*, *pol*, and *env* (Brighty and Rosenberg, 1994; Cochrane *et al.*, 1991; Maldarelli, Martin, and Strebel, 1991; Nasioulas *et al.*, 1994; Rosen *et al.*, 1988; Schwartz, Felber, and Pavlakis, 1992a). In addition, splice sites have been implicated in HIV-1 Rev function (Barksdale and Baker, 1995; Chang and Sharp, 1989; Hammarskjold *et al.*, 1994; Mikaelian *et al.*, 1996; Tan *et al.*, 1996). In this report, the CRS activity of HIV-1 splice sites was investigated through sequential mutagenesis of a proviral clone. In the absence of the majority of HIV-1 intronic sequences, it was observed that Rev-responsiveness was maintained. It is possible that an important component of CRS activity was lost through deletion of these intronic sequences. The intronic sequences may still have a significant impact on CRS activity within the virus.

Unspliced RNAs containing either the first or second *tat* intron exhibited CRS activity. Removal of known splice acceptors did not eliminate the CRS activity present within the RNAs. These RNAs contained either the major 5' or 3' HIV-1 splice donor site. In the absence of any known downstream splice acceptors, these splice donors were still efficiently utilized, in conjunction with cryptic splice acceptors. Deletion of nearly all sequences between the major 5' HIV-1 splice donor and the polyadenylation site appeared to reduce, but did not eliminate, the utilization of the 5' major splice donor. Although deletion of these sequences did not relieve the CRS activity within the unspliced RNA, mutation of the splice donor did relieve CRS activity. Therefore, the major HIV-1 5' splice donor could function as a CRS, even in the absence of

efficient splicing. These results clearly demonstrated that the HIV-1 splice donor sites are intimately involved in the regulation of the cytoplasmic accumulation of viral RNA.

The mutation of the splice donor site within ΔNI also led to an apparent decrease in the total level of RNA. This reduction may be due to a decrease in stability of the RNA caused by the removal of the splice donor. A similar effect on total RNA levels was previously observed when the major 3' HIV-1 splice donor was mutated within a subgenomic *env* construct (Lu *et al.*, 1990); although the effect seen in the previous report was more severe and led to an almost complete inability to detect any RNA produced from the construct.

Results from other groups support the idea that an isolated splice donor might function as a CRS. Naturally occurring inhibitory 5' splice sites within the 3' untranslated region of papillomaviruses were shown to function as CRS in that Rev can counteract their effect (Barksdale and Baker, 1995). The major 5' splice donor of equine infectious anemia virus was demonstrated to function as a CRS (Tan *et al.*, 1996). A ß-globin splice donor was shown to be capable of functioning as a CRS (Chang and Sharp, 1989; Hammarskjold *et al.*, 1994; Mikaelian *et al.*, 1996). The HIV-1 *tev* splice donor, located within the HIV-1 *env* gene, also appears to function as a CRS (Hammarskjold *et al.*, 1994). The results presented in this chapter indicate that authentic HIV-1 splice donors function as CRS to prevent the cytoplasmic accumulation of viral RNAs. The identification of utilized HIV-1 splice donor sites as CRS leaves open the possibility that cryptic HIV-1 splice donor sites might also function as CRS when expressed within the appropriate context.

It seems apparent that the control of HIV-1 RNA production involves a variety of regulatory activities. Splicing serves to downregulate the level of cytoplasmic unspliced and partially spliced RNAs. The splice donors themselves also serve to downregulate the cytoplasmic accumulation of those RNAs in which they are present. The intronic CRS may serve to retain RNA in the nucleus, decrease RNA stability, or decrease RNA cytoplasmic utilization (Brighty and Rosenberg, 1994; Cochrane *et al.*, 1991; Maldarelli, Martin, and Strebel, 1991; Nasioulas *et al.*, 1994; Rosen *et al.*, 1988; Schwartz, Felber, and Pavlakis, 1992a). The combination of each of these regulatory activities may synergize within the virus to prevent viral structural protein production. Rev may serve to overcome all of these levels of negative regulation and permit viral structural protein production.

Chapter 3

Activation of a Rev-Dependent Cryptic Splice Donor in Human Immunodeficiency Virus Type-1

Introduction

By examining isolated sequences, such as splice sites or intronic CRS, it is possible to ascribe functions to these elements. However, unless they are examined together, it is not possible to accurately predict the interplay of splice sites and intronic CRS that are responsible for the Rev regulation of viral expression. Much of the work on CRS and Rev regulation in HIV-1 has been done in disparate systems where relevance of the results becomes an important criticism. Therefore, it is important to establish the biological relevance of the isolated phenotype in the context of a more complete virus. The function of splice sites on Rev-responsiveness in the context of a construct that contains intronic CRS will be examined. It is hoped that this analysis, while more complex, will provide important insight into the relationship between splice sites and intronic CRS and outline possible mechanisms of CRS function.

The HIV-1 protein Rev regulates the expression of viral structural proteins at the level of splicing, transport, stability, and translation (Arrigo and Chen, 1991; Cochrane *et al.*, 1991; D'Agostino *et al.*, 1992; Feinberg *et al.*, 1986; Felber *et al.*, 1989; Hadzopoulou *et al.*, 1989; Hammarskjold *et al.*, 1989; Lawrence *et al.*, 1991; Malim *et al.*, 1988; Schwartz, Felber, and Pavlakis, 1992b; Sodroski *et al.*, 1986) . Rev performs this activity by binding to an RNA binding site termed the Rev-responsive element (RRE) present within the RNAs that it regulates (Hadzopoulou *et al.*, 1989; Malim *et al.*, 1989b; Rosen *et al.*, 1988). Rev may function by alteration of the cellular pathway of the RNA (Fischer *et al.*, 1995). Since Rev is required for protein expression, the Rev dependent RNAs must contain negative sequences, which Rev overcomes for expression (for

review see Dayton 1996) (Dayton, 1996). These sequences should not be present in the Rev-independent *tat*, *rev*, and *nef* genes. These negative sequences have been termed cis-acting repressor sequences, CRS, and have been characterized in a variety of experimental systems. Splice sites have been identified in Rev regulation by several groups (Barksdale and Baker, 1995; Chang and Sharp, 1989; Hammarskjold *et al.*, 1994; Mikaelian *et al.*, 1996; Tan *et al.*, 1996). In addition, the isolated major 5' splice donor of HIV-1 functions as a CRS (Borg, Favaro, and Arrigo, 1997). Introduction of sequences containing the HIV-1 *tat* 3' splice donor or splice acceptor sequences (i.e., those splice sites that join the Tat coding exons) into a construct containing a ß-globin intron conferred Rev-responsiveness to the construct (Chang and Sharp, 1989). Rev-responsiveness could also be conferred by mutation of the ß-globin splice site consensus sequence within this construct. The CRS potential of splice donor sequences has been demonstrated by the Rev-responsiveness of papillomavirus constructs containing a splice donor in the 3' untranslated region (Barksdale and Baker, 1995). The major 5' splice donor of equine infectious anemia virus functions as a CRS within EIAV Gag-producing constructs, although it functions in a location-dependent manner (Tan *et al.*, 1996). Within HIV-1 Env-expressing constructs, splice sites and U1 interaction have been shown to be important for the regulation of Env expression by Rev (Lu *et al.*, 1990). Furthermore, the expression of Env could be made Rev-independent by placing a complete intron upstream of *env* and removing known splice donor sequences (Hammarskjold *et al.*, 1994).

Other sequences within HIV-1, distinct from splice sites, have also been shown to exhibit CRS activity. Recently, a CRS has been identified

within HIV-1 *protease* that functions in an orientation independent fashion to negatively regulate the cytoplasmic accumulation of Rev dependent RNA (Huffman and Arrigo, 1997). The fusion of various fragments of the *gag*, *pol*, and *env* genes to a CAT reporter construct resulted in Rev-responsiveness (Cochrane *et al.*, 1991; Rosen *et al.*, 1988). The CRS within *pol* elicited an orientation-dependent 30-fold negative effect on CAT activity that could be rescued by Rev and the RRE. This CRS is located within HIV-1 sequences that are not known to contain any splice sites. The effect appeared to be at the level of mRNA utilization, in that the subcellular distribution of RNA was not substantially affected. Other studies have shown that fusion of regions of both *gag* and *pol* to reporter constructs demonstrated at least two orientation-dependent CRS that functioned posttranscriptionally (Maldarelli, Martin, and Strebel, 1991). These sequences do not contain known splice sites and were not substantially affected by the addition of splice sites to the constructs. Fusions of *gag* to a *tat*-expression vector were used to analyze CRS activity within *gag* (Schwartz, Felber, and Pavlakis, 1992a). Further research on *gag* sequences showed that a less than 300 base pair fragment was capable of exerting an effect on the total RNA production, which was attributed to differences in RNA stability. Mutational inactivation of this CRS required 28 point mutations, demonstrating the complex nature of this CRS (Schwartz *et al.*, 1992). In the absence of known splice sites, *env* expression was shown to be Rev-dependent in mammalian cells using Env expression vectors (Nasioulas *et al.*, 1994). Localization of CRS within *env* identified many regions capable of exerting modest negative activity, which together exerted more dramatic activity. Deletion of the RRE led to Rev-independent expression of Env from subgenomic *env* constructs,

indicating that a CRS may overlap RRE sequences (Brighty and Rosenberg, 1994; Churchill *et al.*, 1996). The complexity of the interaction between the *gag* CRS and splicing was recently demonstrated (Mikaelian *et al.*, 1996). The *gag* CRS in conjunction with the RRE was capable of conferring Rev-responsiveness on a ß-globin construct. Mutations in splice sites rendered these constructs Rev-independent. In summation, these studies indicate that there are several regions that contain CRS, which determine the viral requirement for Rev.

A potential complication in the study of CRS and the Rev dependence of viral RNAs is the examination of these sequences in the context of other CRS. In order to determine the effect of a mutated splice donor in the context of other viral sequences, the major 5' splice donor was mutated in the context of *gag* sequences to determine the impact the mutation would have on the cytoplasmic accumulation of HIV-1 RNAs. The ratio of cytoplasmic to nuclear HIV-1 RNAs in the presence or absence of Rev was analyzed as an indicator of Rev dependence of the RNA. RT-PCR was utilized to examine the localization of the RNAs produced by the transfected constructs. In this chapter, the effect of the major 5' splice donor, in the context of *gag* sequences, on the Rev-mediated regulation of HIV-1 RNAs was examined.

Materials and Methods

Plasmid Constructions

The plasmids CSF and RoSt-A have been previously described (Campbell, Borg, and Arrigo, 1996; Favaro and Arrigo, 1997). Briefly, CSF is a wild type infectious proviral clone and RoST-A is a Rev mutant in which Rev is deleted through the introduction of a stop codon. RoST-A is also deleted for

a 3600bp *pol* fragment between the Avr II sites which renders the construct noninfectious. RoST-A ΔSD1 was generated from the RoST-A construct using DR mutagenesis (Gatlin *et al.*, 1995) and the following oligonucleotide primers: ΔSD1 5'-AGAACCGCGGACGCCGAAATTTTGACTAGCGGAGGC-3' and ΔSD1A, 5'-ACTACCGCGGGTCGCCGCCCCTCGCCTCTTGC-3'. This results in the deletion of the 5' major splice donor, TGGTGAGT from the RoST-A construct and the insertion of the sequence CCGCGG.

Cell Culture and Transfection

COS cells were maintained in Iscove's medium supplemented with 10% fetal calf serum. Transfections were performed as described previously (Arrigo *et al.*, 1989; Borg, Favaro, and Arrigo, 1997). A total of 50 μg of plasmid DNA was electroporated for each sample. Cotransfections were performed with the indicated construct without or without the addition of 10μg pSVtat and/or 10μg pCMV Rev (Lewis *et al.*, 1990). A filler construct (pcDNA3, Invitrogen) was used to normalize the total amount of DNA to 50μg. A confluent T175cm^2 flask of cells was used for approximately six transfections. RNA was harvested approximately 18 hours post-electroporation.

RNA fractionation and analysis

Cells were fractionated as previously described to obtain two cytoplasmic and a nuclear fraction (Borg, Favaro, and Arrigo, 1997; Favaro and Arrigo, 1997). Isolation of RNA and RT-PCR was performed as previously described (Arrigo and Chen, 1991; Arrigo *et al.*, 1989; Arrigo *et al.*, 1990; Borg, Favaro, and Arrigo, 1997). RT-PCR was performed for 20 cycles of amplification for all RNAs except that of the unspliced RNAs detected with the LA 8 / CSD2 and LA 8/ LA 9 oligonucleotides which

were amplified for 25 cycles. Most of the oligonucleotide primers used for detection have been previously described (Arrigo and Chen, 1991; Arrigo *et al.*, 1989; Arrigo *et al.*, 1990; Borg, Favaro, and Arrigo, 1997). Total RNA was detected with the LDR/αLDR oligonucleotides. Unspliced RNA was detected with several different pairs of oligonucleotides including LA 8 and LA 9, VPR-Xba and LA 23 and LA 8 and CSD2. The antisense oligonucleotide CSD2 hybridizes downstream of the identified cryptic splice donor and its sequence is 'CCCACCTATCCAATTCTC'.

Quantitation of the ratio of cytoplasmic to nuclear RNA was done using a Molecular Dynamics phosphorimager. The cytoplasmic fractions (Cyto 1 and Cyto 2) were added together and compared to the nuclear fraction to determine the ratio of cytoplasmic to nuclear RNA. The level of nuclear RNA was arbitrarily set to a level of 1.

Molecular Cloning and Sequencing of Cryptically Spliced RNA

RT-PCR was performed as described above, on the first cytoplasmic fraction of RNA from the RoST-A ΔSD1 transfection utilizing the LA8/LA23 oligonucleotides. This results in amplification of the spliced RNA. The resulting products were passed over a G50 column (Pharmacia). The DNA products were then ligated into the plasmid pGEM T-easy (Promega). The plasmids were then sequenced to determine the location of the cryptic splice donor. Sequencing of the plasmids was performed using an ABI 377 DNA sequencer for automated sequence analysis. Fluorescent dye-labeled terminators were utilized in conjunction with the T7 primer to obtain the results.

Results

Mutation of the 5' splice donor of HIV-1 does not eliminate splicing to the downstream splice acceptor

Previous studies have examined RNA expression following mutation of the major 5' splice donor of HIV-1 and found that the splice site was responsible for Rev dependence when mutated in isolated splice donor containing constructs (Borg, Favaro, and Arrigo, 1997). One question raised by these experiments was how the 5' major splice donor might interact in the more natural context of the HIV-1 genome. In order to examine this question, the major 5' splice donor was mutated in the Rev-deleted construct RoSt-A to generate RoSt-A ΔSD1 (Figure1). This mutation deletes the consensus 'gt' required for splicing, as well as a previously identified cryptic splice donor immediately downstream (Purcell and Martin, 1993). These constructs are also deleted for *pol* ; to render them noninfectious. The splice acceptors utilized to produce *vif* and *vpr* RNAs are deleted from these constructs. Transient transfections were performed in COS cells to assay for the cytoplasmic accumulation of HIV-1 RNA. The RoSt-A and RoSt-A ΔSD1 constructs were transfected in the presence of Tat and the presence or absence of Rev. A Tat-producing plasmid was cotransfected to provide a source of Tat for the transfections with the deletion of the 5' major splice donor. Since neither RoSt-A nor its derivative RoSt-A ΔSD1 produces Rev, a Rev-producing plasmid or control plasmid was transfected to assess the effect of Rev on the cytoplasmic accumulation of RNA. After transfection, RNA was harvested from the

cells after 18 hours and fractionated into three subcellular fractions. RNA was isolated and analyzed using quantitative RT-PCR procedure.

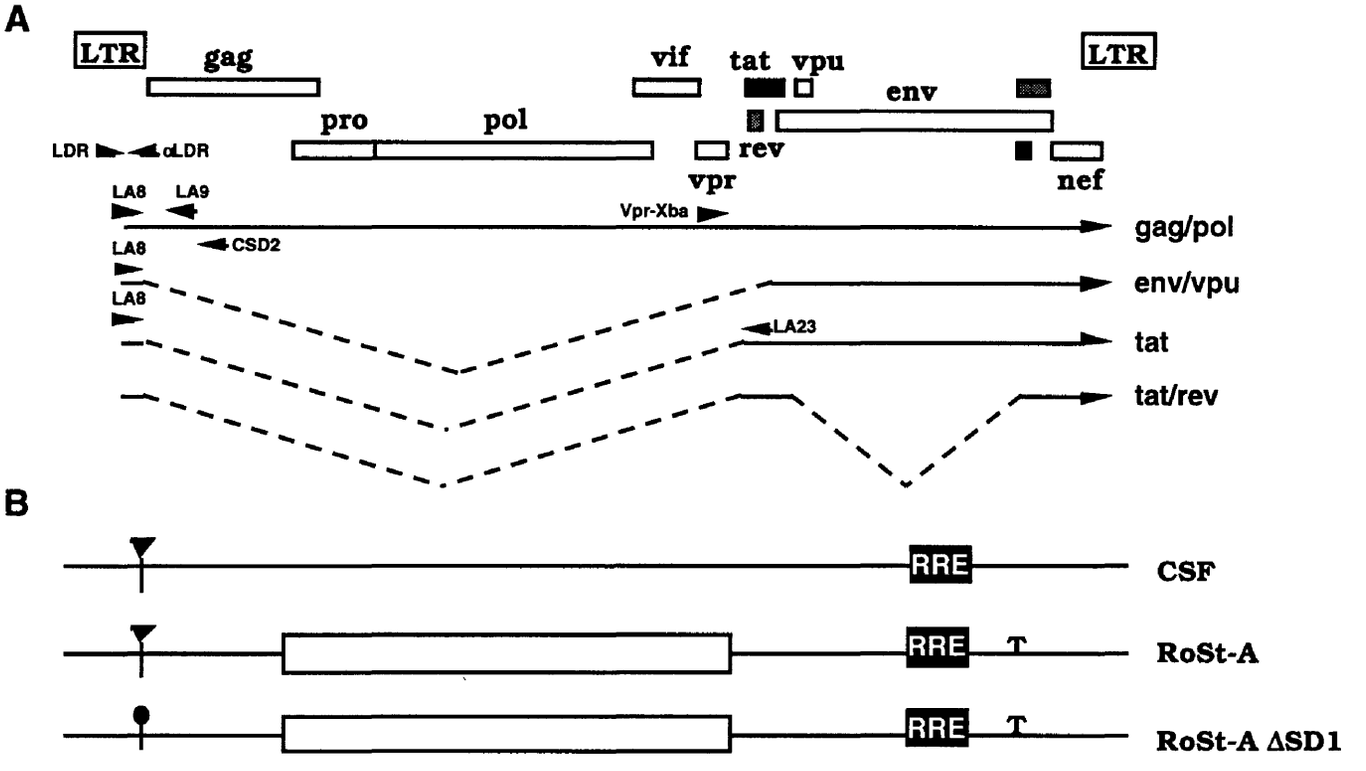


Figure 3.1. Constructs and oligonucleotides used to identify and examine Rev regulation of a cryptic splice donor in HIV-1 gene expression. Schematic diagram of HIV-1 genome, HIV-1 RNAs, and constructs generated. Section **A** depicts the viral genome with hypothetical RNAs produced via splicing. Oligonucleotide primers utilized for detection of various RNAs are depicted as arrows above the RNAs. LDR and αLDR are homologous to sequences in the RNA leader sequence present in all RNAs. LA8 is just upstream and LA9 is just downstream of the major 5' splice donor. CSD2 is just downstream of the cryptic splice donor. LA23 is just downstream of the 5' *tat* splice acceptor. Section **B** depicts the constructs (RoSt-A and RoSt-A ΔSD1) used in these experiments as compared to the wild type strain CSF. The major 5' HIV-1 splice donor is represented as a filled triangle. Deletions are represented as open boxes. The filled circle represents the mutated splice donor site. The location of the RRE in the constructs is shown. The 'T' indicates the addition of a stop codon in RoSt-A and RoSt-A ΔSD1 that blocks Rev expression.

Examination of RNA expression utilizing oligonucleotides that detected all viral RNAs (LDR1/ αLDR1 on Figures 3.1 and 3.2A) indicated no change in distribution with the presence or absence of Rev or with mutation of the 5' major HIV-1 splice donor (see Figure 3.2A). The C:N ratios for the RNA examined with LDR1/ αLDR1 were 2.2 and 1.9 for the RoSt-A (- and + Rev); and 1.6 and 2.5 for the RoSt-A ΔSD1 (- and + Rev) respectively. A ratio of 1:1 indicates that 50% of the RNA was found in the cytoplasmic fractions and 50% was found in the nuclear fraction. The decreased amount of RNA in the RoSt-A ΔSD1 plus Rev cytoplasmic fraction 2 was not reproducible. These results indicate that deletion of the splice donor resulted in little or no change in the total amount of HIV-1 RNA produced by the constructs.

Examination of RNA expression utilizing oligonucleotides that detect the unspliced RNA across the 5' HIV-1 splice donor (LA 8/ LA9 on Figures 3.1 and 3.2A) revealed that the RoSt-A construct expressed cytoplasmic unspliced RNA in a Rev dependent manner. The addition of Rev was required for the cytoplasmic accumulation of unspliced RNA produced by the RoSt-A construct. However, the same RNA produced by the RoSt-A ΔSD1 construct appeared Rev independent after mutation of the splice donor. The C:N ratios for the RNA examined with LA8/ LA 9 were 0.2 and 1.9 for the RoSt-A (- and + Rev); and 1.1 and 1.4 for the RoSt-A ΔSD1 (- and + Rev) respectively. A small shift in the size of this unspliced RNA between the RoSt-A and RoSt-A ΔSD1 constructs was also observed. This is due to the loss of two bases with the splice site mutation. The unspliced RNA at this location was also analyzed utilizing another pair of oligonucleotides (LA 8 and Gag- Nco) and yielded similar results;

the cytoplasmic unspliced RNA produced from RoSt- A was Rev dependent while the cytoplasmic accumulation of unspliced RNA produced from RoSt-A ΔSD1 was independent of Rev (data not shown). These results indicated that the RNA produced by the RoSt-A construct is dependent on Rev for its cytoplasmic accumulation while the cytoplasmic accumulation of RNA produced by the RoSt-A ΔSD1 construct is independent of the presence of Rev.

When an oligonucleotide further downstream from LA 9 (past the cryptic splice donor) is utilized, a Rev-dependent pattern of RNA expression was observed (Figure 3.2). The cytoplasmic accumulation of this product requires the presence of Rev. This oligonucleotide pair, LA 8 and CSD2 (see Figures 3.1 and 3.2A) is similar to LA8 and LA 9, except that CSD2 is about 30bp downstream of LA9. A small shift in the size of this unspliced RNA between the RoSt-A and RoSt-A ΔSD1 constructs was also observed. This is due to the loss of two bases with the splice site mutation. The C:N ratios for the RNA examined with LA8/ CSD2 were 0.13 and 1.9 for the RoSt-A (- and + Rev); and 0.23 and 0.75 for the RoSt-A ΔSD1 (- and + Rev) respectively. These results indicated that the RNA produced by the RoSt-A construct is dependent on Rev for its cytoplasmic accumulation and that the cytoplasmic accumulation of RNA produced by the RoSt-A ΔSD1 construct is also dependent of the presence of Rev.

Analysis of the unspliced RNA with oligonucleotides that span the *tat* splice acceptor (Vpr-Xba and LA 23) revealed a Rev-dependent pattern of RNA expression for both RoSt-A and RoSt-A ΔSD1. The C:N ratios for the RNA examined with Vpr-Xba / LA 23 were 0.38 and 2.1 for the RoSt-A (- and + Rev); and 0.76 and 2.1 for the RoSt-A ΔSD1 (- and + Rev)

respectively. These results indicated that the RNA produced by the RoSt-A construct is dependent on Rev for its cytoplasmic accumulation and that the cytoplasmic accumulation of RNA produced by the RoSt-A ΔSD1 construct is also dependent on the presence of Rev. Since total RNA (LDR1/ αLDR1) ratios are similar in the presence and absence of Rev, but the unspliced (Vpr-Xba and LA 23 and LA 8 and CSD2) are different with respect to Rev-dependence; this indicates that splicing activity must be present in the RoSt-A ΔSD1 construct. This cryptic splice donor must be located between LA 9 and LA 23. The utilization of this splice donor is also dependent on the presence of Rev. This is indicated by the fact that Rev increases the cytoplasmic accumulation of unspliced RNA containing the cryptic splice donor, while in the absence of Rev, cryptic splicing is activated (by examination of total RNA ratios).

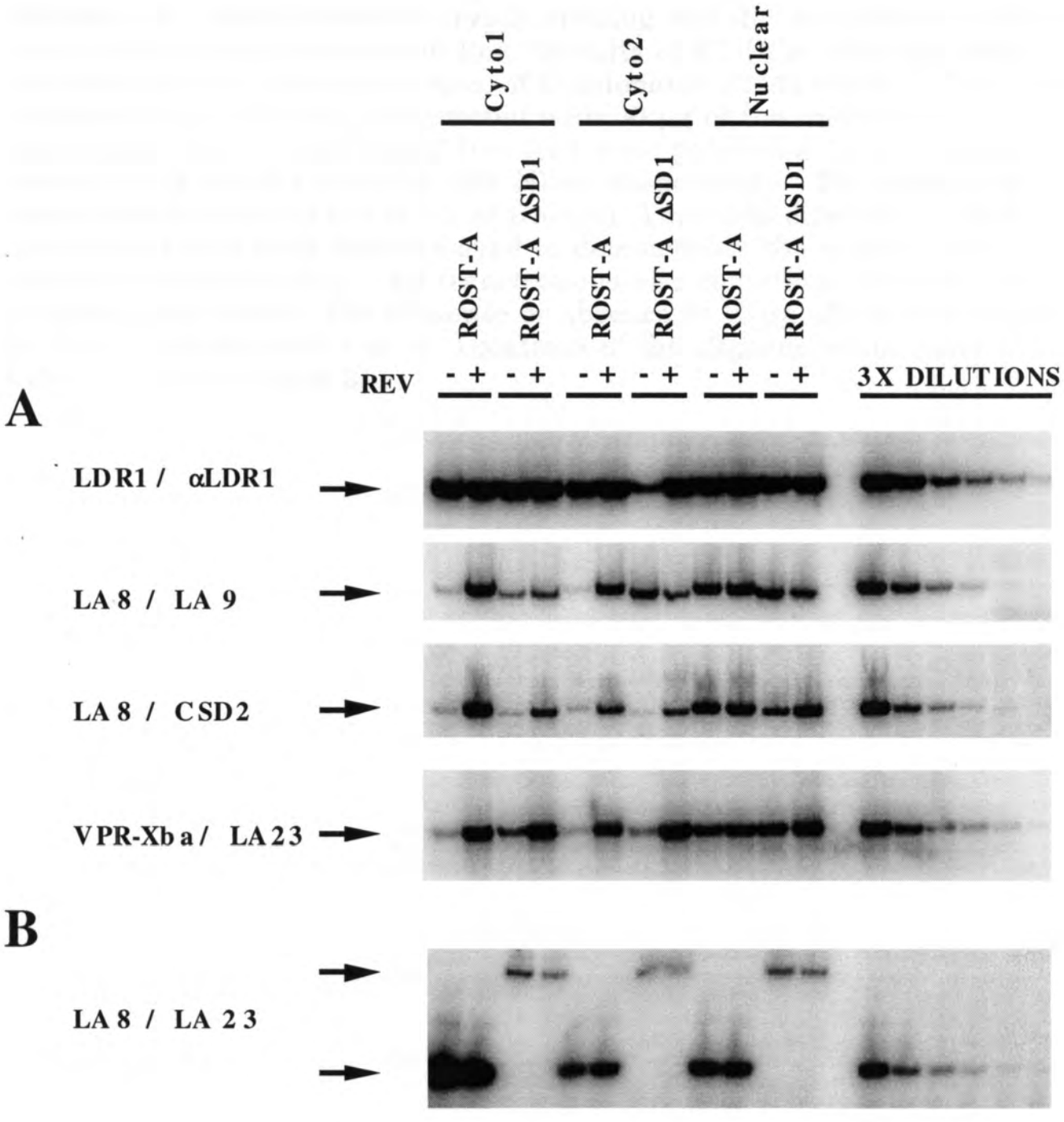

Cyto1
Cyto2
Nuclear
ROST-A
ROST-A ΔSD1
REV
- + - + - + - + - + - +
3X DILUTIONS
A
LDR1 / αLDR1
LA 8 / LA 9
LA 8 / CSD2
VPR-Xba / LA23
B
LA 8 / LA 23

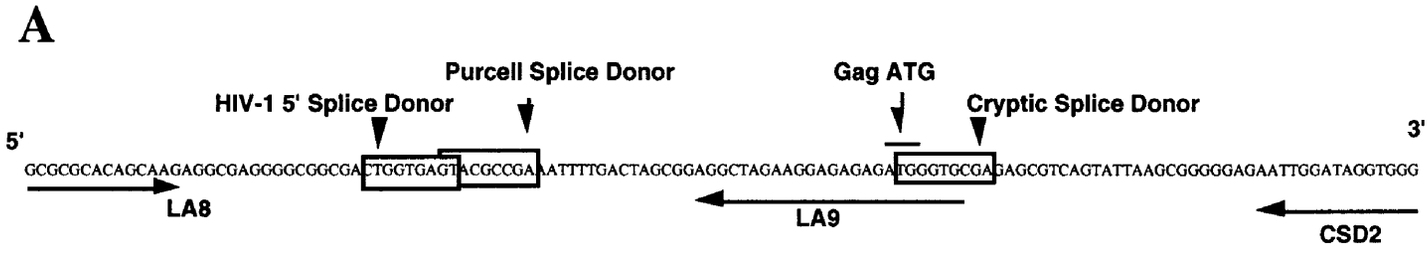
Figure 3.2. Identification of cryptic splicing and Rev dependence of RNAs in the presence and absence of Rev. Results of RT-PCR analysis, using different pairs of oligonucleotides, of fractionated RNAs (Cyto 1, Cyto 2 and Nuclear) from COS cells transfected with 30 μg of the indicated constructs. Cyto 1 and Cyto 2 fractions were generated using a buffer containing 0.05% NP40 and 0.65% NP40, respectively. The remaining pellet was designated the Nuclear fraction. Threefold dilutions of RNA from transfected cells were assayed to demonstrate the quantitative nature of this procedure. All transfections also contained 10 μg of a Tat-producing construct. The presence or absence of 10 μg of a Rev-producing construct is indicated (+ or -). Locations of the oligonucleotide pairs are listed in Figures 1 and 3.

Identification of a Cryptic Splice Donor

The oligonucleotide primer set LA 8 and LA 23 were used to examine the pattern of spliced RNA and to attempt to identify the cryptic splice donor. These oligonucleotides will detect products that are spliced from the 5' major splice donor and downstream of that site to the *tat* splice acceptor (see Figure 3.1). Analysis of RNA with LA 8 and LA 23 indicated a Rev independent pattern of expression of the spliced RNA. The pattern of expression of RNA from both the RoSt-A and the RoSt-A ΔSD1 constructs is similar in that neither is dependent on Rev for the cytoplasmic accumulation of RNA. The spliced RNA produced by the RoSt-A construct is primarily localized in the first cytoplasmic fraction and its cytoplasmic accumulation is largely unaffected by the presence or absence of Rev. The spliced RNA produced by the RoSt-A ΔSD1 construct is also primarily cytoplasmic in its distribution and not greatly affected by the cotransfection of a Rev producing plasmid. Interestingly, there may be a slight decrease in the level of cytoplasmic spliced RNA in the presence of Rev. There is a shift in size between the two species of RNA due to the different location of the splice donor. The cryptic splice donor is located further downstream of LA8 resulting in a larger amplified product.

The spliced RNA from the cytoplasmic 1 fraction of the RoSt-A ΔSD1 minus Rev transfection was amplified using the LA 8 and LA 23 primers (Figure 3.2B). Subsequently, these products were then cloned into the pGEM T-easy vector system for sequencing. The plasmids containing the cloned spliced RNA were then sequenced by an automated ABI 377 sequencer. Sequencing of the cloned RNA with the T7 primer RNA yielded the location of a cryptic splice donor (Figure 3.3A). The sequencing identified the splice donor at 793bp in JRCSF as a 5/9 match to the

consensus sequence (Mount, 1982) This cryptic splice donor lies immediately downstream of the Gag ATG (Figure 3A). This cryptic splice donor was utilized with the *tat* splice acceptor at 5787 bp in JRCSF. The 5/9 match for the identified cryptic site compares to the wild type 5' HIV-1 splice donor which has 8/9 and the 3' major HIV-1 splice donor with 6/9 match to the consensus for splice donors (Figure 3.3B).



B

Consensus Splice Donor Site	C/A A G/ G T A/G A G T
5' HIV-1 Splice Donor Site	C T G/ G T G A G T
3' HIV-1 Splice Donor Site	G C A / G T A A G T
Purcell Cryptic Splice Donor Site	A G A/ G T A C G C
Borg Cryptic Splice Donor Site	T G G/ G T G C G A

Figure 3.3 Sequence of HIV-1 at the 5' major splice donor and comparisons of cryptic splice sites. Section A of the figure depicts the DNA sequence of HIV-1 JRCSF from 711 bp to 837 bp. The locations of the splice donors, oligonucleotides and Gag ATG are indicated. The Purcell splice donor indicates the cryptic splice donor noted by Purcell, et al. Section B of the figure depicts a comparison of the 5' major splice donor, the 3' major splice donor, a previously identified cryptic splice donor (Purcell and Martin, 1993), and the cryptic splice donor identified in this manuscript as compared to the consensus sequence.

Discussion

The major 5' splice donor was analyzed, in the context of *gag* sequences, in order to determine its role on the Rev-mediated regulation of HIV-1 RNAs. RNAs were analyzed by RT-PCR for cytoplasmic accumulation with different combinations of oligonucleotides utilized to assess the localization of different RNAs. The inactivation of the 5' major splice donor could yield one of three possible effects. First, all splicing could be eliminated from the RNA produced by the mutated construct. Second, all other (e.g. the 3' major splice donor) splicing events will remain but splicing in the 5' intronic region to the *tat* splice acceptor will be eliminated. The final possibility is that a cryptic splice donor would be activated in the 5' region and splicing will also occur in the 3' intronic region. The data presented in this chapter indicate that the third possibility is the most probable result of the inactivation of the 5' major splice donor.

Using a pair of oligonucleotides that hybridize to a region common to all viral RNAs, total levels of HIV-1 RNA were detected (LDR1/ αLDR1). A similar distribution of RNA was detected in the cellular fractions. Next, a pair of oligonucleotides that hybridize across the major splice donor was utilized to detect RNA that is unspliced at that location (LA8 and LA 9). In the presence of the wild-type splice donor, this RNA was unspliced and required Rev for cytoplasmic accumulation. One caveat is that splicing may occur downstream of the sequences amplified, hence the PCR product detected is not necessarily entirely unspliced. When the 5' major splice donor was deleted in the RoSt-A ΔSD1 construct this RNA (LA8 and LA 9) is largely independent of Rev for its cytoplasmic accumulation. With activation of a cryptic splice donor downstream, a shift of the 3'

oligonucleotide farther downstream (past the cryptic splice donor) again yields a PCR product that requires Rev for cytoplasmic accumulation (LA 8 and CSD2). Therefore, a cryptic splice site or other CRS element is present downstream of LA 9. Oligonucleotide pairs were also used to examine the unspliced RNA across the *tat* splice acceptor (Vpr-Xba and LA 23). These results when compared to the total RNA levels (LDR1/ αLDR1), which show equal levels of cytoplasmic accumulation, indicate that splicing is still occurring to this splice acceptor. The total ratios of RNA remained unchanged in the presence of Rev, while the unspliced RNA was dependent on Rev for cytoplasmic accumulation; therefore splicing must account for the rest of the total RNA. Analysis of the spliced RNA (LA 8 and LA 23) indicated that the spliced products were independent of Rev and illustrated that a cryptic splice donor was splicing to the *tat* splice acceptor site.

The sequence of this site was identified and compared to the consensus sequence. It is notable that such a poor match (5/9) to the consensus sequence was utilized as a splice site and regulated by Rev. However, it is only one basepair less of a match than the 3' major splice donor of HIV-1. A 5/9 match cryptic splice donor has also been reported previously in the literature (Purcell and Martin, 1993) (see Figure 3.3). Presence of this cryptic splice is especially interesting in light of the fact that there are a large number of possible cryptic sites when *gag* sequences were analyzed for consensus matches. Some of these sites are even better matches than this identified site. There are probably multiple cryptic sites activated upon mutation of the 5' major splice donor. However, no other cryptic splicing events were detected by RT-PCR analysis utilizing a large number of different oligonucleotide pairs throughout *gag*.

The fact that Rev regulation requires splice site recognition has been noted previously (Hammarskjold *et al.*, 1994; Lu *et al.*, 1990). Another study has reported that a cryptic splice donor is activated upon mutation of the wild-type splice donor (Purcell and Martin, 1993). The data presented here demonstrate activation of a cryptic splice donor with the inactivation of the 5' major splice donor and a previously identified cryptic splice donor. It was not unexpected that the unspliced RNA expressed from the RoSt-A ΔSD1 construct might remain Rev dependent due to the retention of numerous CRS that have been previously described (for review see Dayton) (Dayton, 1996) and that splicing might be present upon activation of cryptic splice donors. However, it is remarkable that the utilization of this splice donor would be dependent on Rev. In retrospect, this fact should not be surprising, as others have reported that suboptimal splice sites appear to be required for Rev regulation of splicing (Katz, 1990; O'Reilly, McNally, andBeemon, 1995). This is perhaps because splicing must be inefficient enough for Rev to differentially regulate the splicing process yet still able to have those sites interact with cellular splicing factors. The life history of HIV requires that it adapt to its ever changing genome. Consequently, the observation that a cryptic splice site immediately downstream of the Gag ATG is regulated by Rev further confirms the plastic nature of the HIV genome. These studies also present the possibility that cryptic splice donors may function as CRS in concert with other identified CRS to determine the Rev-dependence of viral RNAs.

Chapter 4

Perspectives

The question of how splice donors affect the cytoplasmic accumulation of HIV-1 RNAs is part of a much larger discussion of how cells control the metabolism and expression of RNA. Many of the difficulties and problems in understanding HIV-1 gene regulation result from our lack of molecular understanding of how cells intrinsically control these processes. Viruses provide an invaluable tool to probe the cellular machinery and yield remarkable insight about critical control points in the complex machinations of gene expression. The findings presented here add to the understanding of how viruses regulate RNA expression utilizing cellular processes to control their gene expression.

The finding that an isolated splice donor functions as a CRS raises several questions. One is, whether splice donors function as a CRS through splicing or by binding a cellular factor and having an effect distinct from the splicing reaction. It is obvious that splicing decreases the amount of unspliced RNA in the cytoplasm. However, it is not clear that splice donors function as CRS merely through splicing of the RNA. Indeed the data presented in Chapter 2 support the interesting conclusion that an isolated splice donor in the presence of minimal splicing functions as a CRS. Ignoring for a moment the minimal splicing detected in these experiments, the findings show that a splice donor can function to negatively regulate the cytoplasmic accumulation of unspliced RNAs. The molecular event probably involves the binding of cellular factors such as U1 which have been shown to be important in CRS function (Lu *et al.*, 1990). Binding of U1 and other snRNPs must direct the RNA, in the absence of splicing to a degradative pathway. This must be true since the

RNA does not accumulate in the cytoplasm nor is it significantly spliced before transport from the nucleus.

However, minimal splicing was detected in the isolated splice donor constructs and the possible mechanisms of this splicing are unclear with the removal of all known splice acceptor sequences and indeed almost all sequences downstream of the splice donor. Possible mechanisms for this might include trans or inverse splicing, rare but well described splicing events (Moore, Query, and Sharp, 1993). Alternatively, readthrough of the polyadenylation signal may be responsible for generating a larger RNA that would contain possible 3' cryptic splice acceptor sequences and allow for splicing. Indeed, readthrough of the polyadenyaltion signal in the 5' LTR occurs in HIV-1 gene expression, providing a precedent for this mechanism. In order to investigate these hypotheses we performed 3' rapid amplification of cDNA ends (RACE) to investigate the sequence of the spliced RNA but the results were inconclusive.

Another interesting question raised by the findings in Chapter 2 is the difference in function between the two major splice donors of HIV-1. The ratios indicate that the second major splice donor is much less of a CRS than the 5' splice donor. It also is 3bp less of a match to the consensus sequence. It might seem probable that the degree of consensus sequence might function to determine the "strength" of a CRS. This might be true because an increase in binding of cellular factors to the CRS could increase the negative regulatory function of the CRS. The binding of splicing factors RNAs has been shown to retain RNAs in the nucleus (Stutz and Rosbash, 1994). It is important to remember that splicing must be regulated for the production of the different classes of viral RNA. In order for this regulation to occur, splicing of viral mRNA must be

inherently inefficient. If splice sites yield efficient splicing, then the reaction would be completed and it would be impossible to rescue unspliced RNA. The finding that an isolated splice donor can regulate the fate of an RNA, without splicing is novel and interesting. How does this relate back to wild-type viral expression? The next set of experiments was designed to investigate that question.

Experiments in Chapter 3 address questions of splice donor function in the context of *gag* sequences. It is difficult to assess the impact of a particular CRS, such as the 5' major splice donor of HIV-1, in the context of other CRS. Mutation of one CRS does not eliminate negative regulation of the RNA because of the redundancy of these elements. These experiments identified a cryptic splice donor that is Rev-responsive. These findings raise several questions including- why do cryptic sites appear upon inactivation of wild-type sites? The answer is not readily apparent but several hypotheses are possible. One explanation for the presence of multiple cryptic sites is the presence of Rev. Since Rev is present during the expression of unspliced RNA to rescue the RNA there is no negative selection for the presence of cryptic elements. Indeed, with significant mutation rates in HIV-1 there maybe some selective pressure to maintain cryptic sequences so that the unspliced RNAs might remain Rev dependent even with a mutation of the wild-type splice donor. The reason for the redundancy of CRS function may never be known but it is curious to speculate on possible reasons for this feature of HIV-1 gene regulation.

Another interesting finding from Chapter 3 was the location of the cryptic splice donor that was identified. It is the next possible site downstream of the previous cryptic splice donor (Purcell and Martin, 1993) which is immediately downstream of the wild type site. It is perhaps

significant that cryptic site selection appears to proceed 5' to 3' through the RNA. Also supporting this is the observation that several better possible cryptic sites exist in *gag* but do not appear to be utilized for cryptic splicing.

These data suggest that all CRS might be splice sites or splice site-binding elements. In support of this are the large number of possible cryptic sites that exist throughout HIV-1. Perhaps the intronic CRS discussed previously are merely sequences (or structures) that bind snRNPs and retain the unspliced RNA in the nucleus. Many of the identified CRS have been examined for splice sites and found to contain none but this does not rule out this mechanism of function. A consensus match may not be required for the binding of a particular factor that in association with a number of other proteins retain the unspliced RNAs in the nucleus. Such a mechanism would be extremely difficult to test, but might be a likely explanation of CRS function.

One piece of evidence that might help evaluate this hypothesis is the finding of a CRS in protease that functions in an orientation independent manner and negatively regulated the cytoplasmic accumulation of RNA (Huffman and Arrigo, 1997). These sequences were examined for splicing events and none were detected. However, sequence analysis suggests that possible cryptic sites might exist. This may indicate that splicing was not present, but could also mean that snRNPs could still be bound to the RNA that retain the CRS containing RNA in the nucleus. Another note is that this sequence appears much weaker than the splice site CRS, which in addition to the orientation effect, is in support of a different mechanism of function. In light of this and other evidence it is difficult to assess the

nature of intronic CRS and whether they are cryptic splice sites or cryptic splice site binding elements.

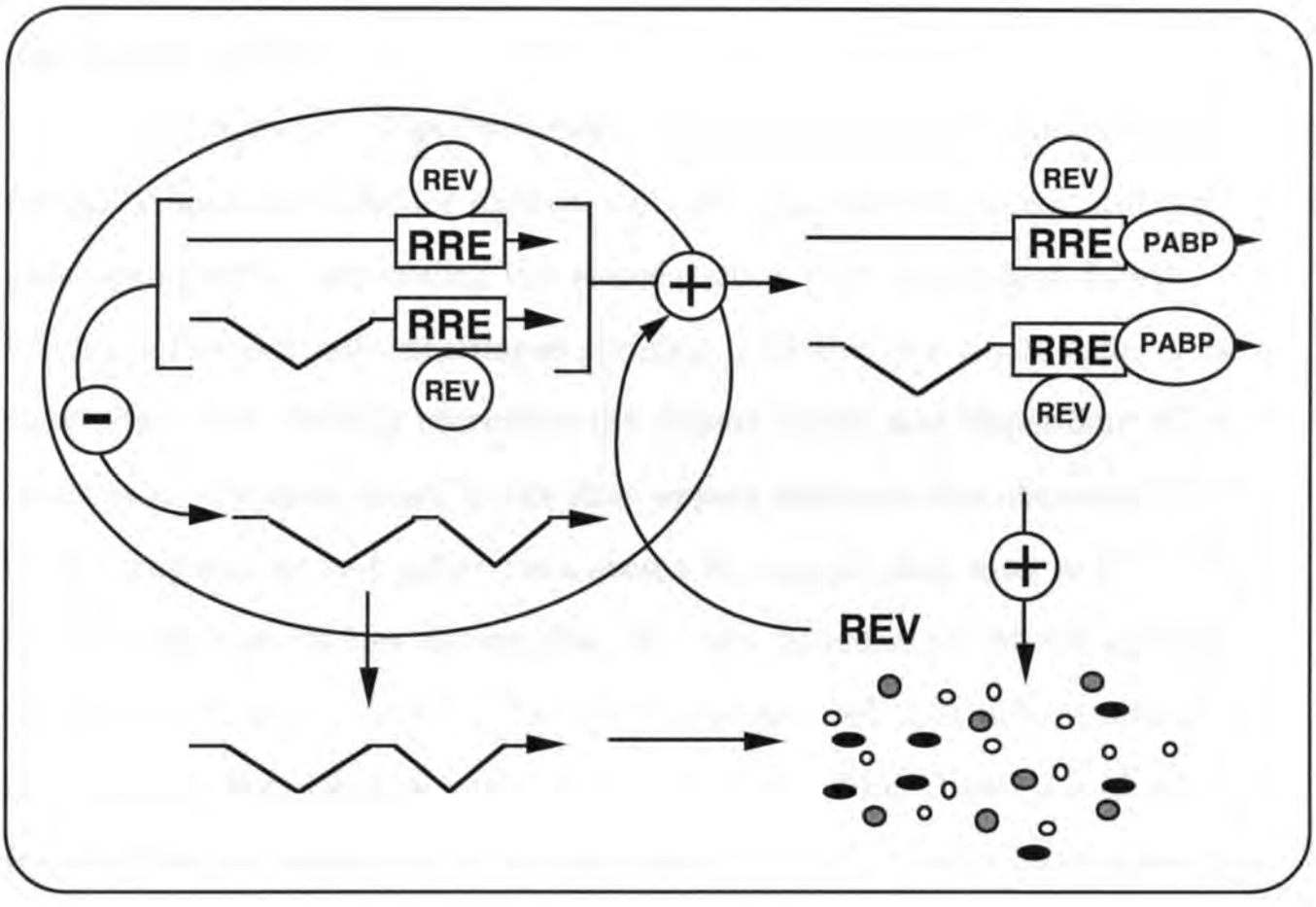


Figure 4.1 This is a hypothetical model of Rev function and its interaction with viral RNAs. The outside boundary is a cell membrane and the large oval structure represents the nucleus. The binding of Rev to various RNAs is depicted as are some of the feedback loops involved in Rev regulation of viral RNA expression.

In light of these findings, a potential model of Rev function is shown in Figure 4.1. This figure shows the expression of Rev from the fully spliced RNA. This RNA is translated to produce Rev, which is localized to the nucleus via its NLS. Once in the nucleus Rev binds to the Rev response element located in Env. The binding of Rev then allows for rescue of the Rev-dependent RNA. Previous work has defined the role of the Rev response element (RRE) in HIV-1 gene regulation. The location of

the RRE has an important impact on RNA expression (Campbell, Borg, and Arrigo, 1996).

The binding of Rev "rescues" the Rev-dependent unspliced and partially spliced RNAs for expression. Rev may also decrease its own expression while increasing the expression of Rev-dependent RNAs. Whether Rev actually decreases splicing, and thereby production of Rev, or whether Rev merely increases the export of the Rev dependent RNAs is uncertain. Popular belief holds that export explains the observed decrease in Rev expression but other data seems to suggest this may not be entirely true. Recent work has shown that Rev can function to inhibit splicing (Favaro and Arrigo, 1997). This work also showed that differences in cell type change the cellular requirement for Rev. While Cos cells absolutely require Rev for cytoplasmic accumulation of RNA, T cells have a low level of cytoplasmic accumulation in the absence of Rev (Arrigo and Chen, 1991; Favaro and Arrigo, 1997). Rev then facilitates the cytoplasmic accumulation of unspliced and partially spliced RNAs but also has further effects on the fate of the RNA.

In the cytoplasm, Rev may also have critical effects on the regulation of RNA expression. Rev may have an important effect on translation of the Rev-dependent RNAs. This is supported by the observation that Rev is not absolutely required for cytoplasmic accumulation of Rev-dependent RNAs in the absence of Rev in lymphoid cells, whereas no expression of proteins occurs in the absence of Rev (Arrigo and Chen, 1991; Favaro and Arrigo, 1997). This level of control may involve association of the RNA with poly-A binding protein (PAB1). PAB1 binds to polyadenylated RNAs and facilitates 60s ribosomal subunit binding. In the absence of Rev, the Rev-dependent RNAs are defective in

their association with PAB1 (Campbell *et al.*, 1994). This level of regulation may explain the lack of translation of these RNAs.

Rev then has an impressive series of functions for a 19Kd protein. After binding to the RRE, Rev may function in splicing, export, stability and translation of the Rev-dependent RNAs as well as regulating its own production. How does Rev accomplish all of this? One theory is that Rev merely directs the RNA to different cellular pathways. By associating with a number of different cellular factors, it might influence the processing by the cellular machinery. In so doing, Rev appears to have many effects. This is especially true when you consider the many different systems used to analyze Rev function. As systems are different, the primary effect of Rev on the fate of the RNA may appear different. Interestingly, the ideas that were originally controversial are now becoming more accepted. The dominant export paradigm of Rev function is still questioned. This is not to say that Rev does not regulate export, merely that other effects are also critical functions of Rev in the regulation of RNA expression.

In the regulation of gene expression, splice sites are critical features that direct RNA metabolism. HIV-1 utilizes the features of this processing to variably regulate the production of different RNAs from a single primary transcript. By using splice sites to control the expression of unspliced RNA, HIV-1 controls the temporal expression of the viral genes. The story of HIV-1 gene regulation is a complex tapestry of interwoven regulatory systems that not only provide important clues to understanding cellular gene expression but also contains the knowledge required to treat and prevent the Acquired Immune Deficiency Syndrome.

References

Adams, M. D., Rudner, D. Z., and Rio, D. C. (1996). Biochemistry and Regulation of pre-mRNA Splicing. *Current Opinion in Cell Biology* **8,** 331-339.

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J. (1994). "Molecular Biology of the Cell,." Third Edition ed. Garland Publishing, Inc, New York.

Amendt, B. A., Si, Z. H., and Stoltzfus, C. M. (1995). Presence of exon splicing silencers within human immunodeficiency virus type 1 tat exon 2 and tat-rev exon 3: evidence for inhibition mediated by cellular factors. *Mol Cell Biol* **15**(8), 4606-15.

Arrigo, S., and Beemon, K. (1988). Regulation of Rous sarcoma virus RNA splicing and stability. *Mol Cell Biol* **8**(11), 4858-67.

Arrigo, S. J., and Chen, I. S. Y. (1991). Rev Is Necessary for Translation But Not Cytoplasmic Accumulation of HIV-1 vif, vpr, and env/vpu-2 RNAs. *Genes & Development* **5**(5), 808-819.

Arrigo, S. J., Heaphy, S., and Haines, J. K. (1992). In vivo binding of wild-type and mutant human immunodeficiency virus type 1 Rev proteins: implications for function. *J. Virol.* **66**(9), 5569-5575.

Arrigo, S. J., Weitsman, S., Rosenblatt, J. D., and Chen, I. S. (1989). Analysis of rev gene function on human immunodeficiency virus type 1 replication in lymphoid cells by using a quantitative polymerase chain reaction method. *J. Virol.* **63**(11), 4875-81.

Arrigo, S. J., Weitsman, S., Zack, J. A., and Chen, I. S. (1990). Characterization and expression of novel singly spliced RNA species of human immunodeficiency virus type 1. *J Virol* **64**(9), 4585-8.

Barksdale, S. K., and Baker, C. C. (1995). The human immunodeficiency virus type 1 Rev protein and the Rev-responsive element counteract the effect of an inhibitory 5' splice site in a 3' untranslated region. *Mol Cell Biol* **15**(6), 2962-71.

Berger, J., Aepinus, C., Dobrovnik, M., Fleckenstein, B., Hauber, J., and Bohnlein, E. (1991). Mutational analysis of functional domains in the HIV-1 Rev trans-regulatory protein. *Virology* **183**(2), 630-635.

Bogerd, H. P., Fridell, R. A., Madore, S., and Cullen, B. R. (1995). Identification of a novel cellular cofactor for the Rev/Rex class of retroviral regulatory proteins. *Cell* **82**(3), 485-94.

Borg, K. T., Favaro, J. P., and Arrigo, S. J. (1997). Involvement of human immunodeficiency virus type-1 splice sites in the cytoplasmic accumulation of viral RNAs. *Virology* **236**(1), 95-103.

Brighty, D. W., and Rosenberg, M. (1994). A cis-acting repressive sequence that overlaps the Rev-responsive element of human immunodeficiency virus type 1 regulates nuclear retention of env mRNAs independently of known splice signals. *Proc Natl Acad Sci U S A* **91**(18), 8314-8.

Campbell, L. H., Borg, K. T., and Arrigo, S. J. (1996). Differential effects of intronic and exonic locations of the human immunodeficiency virus type 1 Rev-responsive element. *Virology* **219**(2), 423-31.

Campbell, L. H., Borg, K. T., Haines, J. K., Moon, R. T., Schoenberg, D. R., and Arrigo, S. J. (1994). Human immunodeficiency virus type 1 Rev is required in vivo for binding of poly(A)-binding protein to Rev-dependent RNAs. *J Virol* **68**(9), 5433-8.

Chang, D. D., and Sharp, P. A. (1989). Regulation by HIV Rev depends upon recognition of splice sites. *Cell* **59**(5), 789-95.

Churchill, M. J., Moore, J. L., Rosenberg, M., and Brighty, D. W. (1996). The rev-responsive element negatively regulates human immunodeficiency virus type 1 env mRNA expression in primate cells. *J Virol* **70**(9), 5786-90.

Cochrane, A. W., Jones, K. S., Beidas, S., Dillon, P. J., Skalka, A. M., and Rosen, C. A. (1991). Identification and characterization of intragenic sequences which repress human immunodeficiency virus structural gene expression. *J Virol* **65**(10), 5305-5313.

D'Agostino, D. M., Felber, B. K., Harrison, J. E., and Pavlakis, G. N. (1992). The Rev protein of human immunodeficiency virus type 1 promotes

polysomal association and translation of gag/pol and vpu/env mRNAs. *Mol Cell Biol* **12**(3), 1375-86.

Dayton, A. I. (1996). The Rev axis of HIV-1 and its associated host cofactors: a viral window onto the workings of eukaryotic posttranscriptional RNA processing. *J. Biomed. Sci.* **3**, 69-77.

Favaro, J. P., and Arrigo, S. J. (1997). Characterization of Rev function using subgenomic and genomic constructs in T and COS cells. *Virology* **228**, 29-38.

Feinberg, M. B., Jarrett, R. F., Aldovini, A., Gallo, R. C., and Wong, S. F. (1986). HTLV-III expression and production involve complex regulation at the levels of splicing and translation of viral RNA. *Cell* **46**(6), 807-17.

Felber, B. K., Hadzopoulou, C. M., Cladaras, C., Copeland, T., and Pavlakis, G. N. (1989). rev protein of human immunodeficiency virus type 1 affects the stability and transport of the viral mRNA. *Proc Natl Acad Sci U S A* **86**(5), 1495-9.

Fischer, U., Huber, J., Boelens, W. C., Mattaj, I. W., and Luhrmann, R. (1995). The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell* **82**(3), 475-83.

Fischer, U., Meyer, S., Teufel, M., Heckel, C., Luhrmann, R., and Rautmann, G. (1994). Evidence that HIV-1 Rev directly promotes the nuclear export of unspliced RNA. *Embo J* **13**(17), 4105-12.

Fritz, C. C., Zapp, M. L., and Green, M. R. (1995). A human nucleoporin-like protein that specifically interacts with HIV Rev. *Nature* **376**(6540), 530-3.

Garrett, E. D., and Cullen, B. R. (1992). Comparative analysis of Rev function in human immunodeficiency virus type-1 and type-2. *Journal of Virology* **66**(7), 4288-4294.

Gatlin, J., Campbell, L. H., Schmidt, M. G., and Arrigo, S. J. (1995). Direct-rapid (DR) mutagenesis of large plasmids using PCR. *Biotechniques* **19**(4), 559-64.

Hadzopoulou, C. M., Felber, B. K., Cladaras, C., Athanassopoulos, A., Tse, A., and Pavlakis, G. N. (1989). The rev (trs/art) protein of human immunodeficiency virus type 1 affects viral mRNA and protein expression via a cis-acting sequence in the env region. *J Virol* **63**(3), 1265-74.

Hammarskjold, M. L., Heimer, J., Hammarskjold, B., Sangwan, I., Albert, L., and Rekosh, D. (1989). Regulation of human immunodeficiency virus env expression by the rev gene product. *J Virol* **63**(5), 1959-66.

Hammarskjold, M. L., Li, H., Rekosh, D., and Prasad, S. (1994). Human immunodeficiency virus env expression becomes Rev-independent if the env region is not defined as an intron. *J Virol* **68**(2), 951-8.

Huffman, K. M., and Arrigo, S. J. (1997). Identification of cis-acting repressor activity within human immunodeficiency virus type-1 protease. *Virology* **234**(2), 253-260.

Katahira, J., Ishizaki, T., Sakai, H., Adachi, A., Yamamoto, K., and Shida, H. (1995). Effects of translation initiation factor eIF-5A on the functioning of human T-cell leukemia virus type I Rex and human immunodeficiency virus Rev inhibited trans dominantly by a Rex mutant deficient in RNA binding. *J Virol* **69**(5), 3125-33.

Katz, R. A. (1990). Control of Retroviral RNA Splicing through Maintenance of Suboptimal Processing Signals. *Molecular and Cellular Biology* **Feb,** 696-704.

Kjems, J., Brown, M., Chang, D. D., and Sharp, P. A. (1991). Structural analysis of the interaction between the human immunodeficiency virus Rev protein and the Rev response element. *Proc Natl Acad Sci U S A* **88**(3), 683-7.

Kjems, J., Frankel, A. D., and Sharp, P. A. (1991). Specific regulation of messenger RNA splicing invitro by a peptide from HIV-1 Rev. *Cell* **67**(1), 169-178.

Lawrence, J. B., Cochrane, A. W., Johnson, C. V., Perkins, A., and Rosen, C. A. (1991). The HIV-1 Rev protein: a model system for coupled RNA transport and translation. *New Biol* **3**(12), 1220-32.

Lewis, N., Williams, J., Rekosh, D., and Hammarskjold, M. L. (1990). Identification of a cis-acting element in human immunodeficiency virus type 2 (HIV-2) that is responsive to the HIV-1 rev and human T-cell leukemia virus types I and II rex proteins. *J Virol* **64**(4), 1690-7.

Lu, X. B., Heimer, J., Rekosh, D., and Hammarskjold, M. L. (1990). U1 small nuclear RNA plays a direct role in the formation of a rev-regulated human immunodeficiency virus env mRNA that remains unspliced. *Proc Natl Acad Sci U S A* **87**(19), 7598-602.

Maldarelli, F., Martin, M. A., and Strebel, K. (1991). Identification of posttranscriptionally active inhibitory sequences in human immunodeficiency virus type-1 RNA - novel level of gene regulation. *J Virol* **65**(11), 5732-5743.

Malim, M. H., Bohnlein, S., Hauber, J., and Cullen, B. R. (1989a). Functional dissection of the HIV-1 Rev trans-activator--derivation of a trans-dominant repressor of Rev function. *Cell* **58**(1), 205-14.

Malim, M. H., Hauber, J., Fenrick, R., and Cullen, B. R. (1988). Immunodeficiency virus rev trans-activator modulates the expression of the viral regulatory genes. *Nature* **335**(6186), 181-3.

Malim, M. H., Hauber, J., Le, S. Y., Maizel, J. V., and Cullen, B. R. (1989b). The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature* **338**(6212), 254-7.

Mann, D. A., Mikaelian, I., Zemmel, R. W., Green, S. M., Lowe, A. D., Kimura, T., Singh, M., Butler, P. J., Gait, M. J., and Karn, J. (1994). A molecular rheostat. Co-operative rev binding to stem I of the rev-response element modulates human immunodeficiency virus type-1 late gene expression. *J Mol Biol* **241**(2), 193-207.

Mikaelian, I., Krieg, M., Gait, M. J., and Karn, J. (1996). Interactions of INS (CRS) elements and the splicing machinery regulate the production of Rev-responsive RNAs. *J. Mol. Biol.* **257**, 246-264.

Moore, M. J., Query, C. C., and Sharp, P. A. (1993). Splicing of Precursors to mRNA by the Spliceosome. *In* "The RNA World", pp. 303-357. Cold Spring Harbor Laboratory Press.

Mount, S. M. (1982). A catalogue of splice junction sequences. *Nucleic Acids Res* **10**(2), 459-72.

Nasioulas, G., Zolotukhin, A. S., Tabernero, C., Solomin, L., Cunningham, C. P., Pavlakis, G. N., and Felber, B. K. (1994). Elements distinct from human immunodeficiency virus type 1 splice sites are responsible for the Rev dependence of env mRNA. *J Virol* **68**(5), 2986-93.

O'Reilly, M. M., McNally, M. T., and Beemon, K. L. (1995). Two strong 5' splice sites and competing, suboptimal 3' splice sites involved in alternative splicing of human immunodeficiency virus type 1 RNA. *Virology* **213**(2), 373-85.

Olsen, H. S., Cochrane, A. W., Dillon, P. J., Nalin, C. M., and Rosen, C. A. (1990). Interaction of the human immunodeficiency virus type 1 Rev protein with a structured region in env mRNA is dependent on multimer formation mediated through a basic stretch of amino acids. *Genes Dev* **4**(8), 1357-64.

Purcell, D. F., and Martin, M. A. (1993). Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. *J Virol* **67**(11), 6365-78.

Rosen, C. A., Terwilliger, E., Dayton, A., Sodroski, J. G., and Haseltine, W. A. (1988). Intragenic cis-acting art gene-responsive sequences of the human immunodeficiency virus. *Proc Natl Acad Sci U S A* **85**(7), 2071-5.

Ruhl, M., Himmelspach, M., Bahr, G. M., Hammerschmid, F., Jaksche, H., Wolff, B., Aschauer, H., Farrington, G. K., Probst, H., Bevec, D., and et, a. l. (1993). Eukaryotic initiation factor 5A is a cellular target of the human immunodeficiency virus type 1 Rev activation domain mediating trans-activation. *J Cell Biol.*

Sachs, A., and Wahle, E. (1993). Poly(A) tail metabolism and function in eucaryotes. *J Biol Chem* **268**(31), 22955-8.

Sachs, A. B., Bond, M. W., and Kornberg, R. D. (1986). A single gene from yeast for both nuclear and cytoplasmic polyadenylate-binding proteins: Domain structure and expression. *Cell* **45**, 827-835.

Sachs, A. B., and Davis, R. W. (1989). The poly(A) binding protein is required for poly(A) shortening and 60S ribosomal subunit-dependent translation initiation. *Cell* **58**(5), 857-67.

Sachs, A. B., and Deardorff, J. A. (1992). Translation initiation requires the PAB-dependent poly(A) ribonuclease in yeast. *Cell* **70**(6), 961-73.

Schwartz, S., Campbell, M., Nasioulas, G., Harrison, J., Felber, B. K., and Pavlakis, G. N. (1992). Mutational inactivation of an inhibitory sequence in human immunodeficiency virus type 1 results in Rev-independent gag expression. *J Virol* **66**(12), 7176-82.

Schwartz, S., Felber, B. K., and Pavlakis, G. N. (1992a). Distinct RNA sequences in the gag region of human immunodeficiency virus type-1 decrease RNA stability and inhibit expression in the absence of Rev protein. *J Virol* **66**(1), 150-159.

Schwartz, S., Felber, B. K., and Pavlakis, G. N. (1992b). Mechanism of translation of monocistronic and multicistronic human immunodeficiency virus type-1 mRNAs. *Mol Cell Biol* **12**(1), 207-219.

Sodroski, J., Goh, W. C., Rosen, C., Dayton, A., Terwilliger, E., and Haseltine, W. (1986). A second post-transcriptional trans-activator gene required for HTLV-III replication. *Nature* **321**(6068), 412-7.

Stutz, F., Neville, M., and Rosbash, M. (1995). Identification of a novel nuclear pore-associated protein as a functional target of the HIV-1 Rev protein in yeast. *Cell* **82**(3), 495-506.

Stutz, F., and Rosbash, M. (1994). A functional interaction between Rev and yeast pre-mRNA is related to splicing complex formation. *Embo J* **13**(17), 4096-104.

Tan, W., Schalling, M., Zhao, C., Luukkonen, M., Nilsson, M., Fenyo, E. M., Pavlakis, G. N., and Schwartz, S. (1996). Inhibitory activity of the equine infectious anemia virus major 5' splice site in the absence of Rev. *J. Virol.* **70**(6), 3645-3658.

Ullman, K. S., Powers, M. A., and Forbes, D. J. (1997). Nuclear Export Receptors: From Importin to Exportin. *Cell* **90**, 967-970.

Biographical sketch

Keith Thomas Borg

Education

Macalester College, 1990, Bachelor of Arts,
Graduated with Honors in Biology.
Minor- Japanese Studies
Current Program- Medical Scientist Training Program (MD/PhD) at the Medical University of South Carolina

Awards

National Student Research Forum 1996, 2nd place poster.
MUSC Research Day 1996, 2nd place poster.
MUSC Research Day 1995, 1st place oral presentation

Abstracts presented-

Keith T. Borg, Justin Favaro, and Salvatore J. Arrigo. Involvement of HIV-1 Splice Sites in the Cytoplasmic Accumulation of Viral RNAs. National MD/Ph.D. Student Conference, Aspen July 1997.

Keith T. Borg, Justin Favaro, and Salvatore J. Arrigo. Involvement of HIV-1 Splice Sites in the Cytoplasmic Accumulation of Viral RNAs. Cold Spring Harbor Retrovirus Meeting, May 1997.

Keith Borg and Salvatore Arrigo. Cis Repressor Sequences and the Regulation of HIV-1 Gene Expression. National Student Research Forum, April 1996, Galveston, Texas.

Lia Campbell, Keith Borg, and Salvatore Arrigo. HIV-1 Rev Relieves "Masking" of the Poly(A) Tail; 2nd National Conference on Human Retroviruses and Related Infections. January 1995, Washington, D.C.

Manuscripts

Keith T. Borg, Justin P. Favaro, Salvatore Arrigo and Michael Schmidt. Activation of a Cryptic Splice Donor in HIV-1. In Press Journal of Biomedical Sciences.

Justin P. Favaro, Keith T. Borg, Salvatore Arrigo and Michael Schmidt. Effect of Rev on the Intranuclear Localization of HIV-1 Unspliced RNA. submitted to Virology.

Keith T. Borg, Justin Favaro, and Salvatore J. Arrigo. Involvement of HIV-1 Splice Sites in the Cytoplasmic Accumulation of Viral RNAs. Virology. 236, 95-103. 1997.

Keith T. Borg, William Burgess, Louis Terracio, Thomas K. Borg. Expression of Metalloproteases by Cardiac Myocytes and Fibroblasts In Vitro. Cardiovascular Pathology. 6:5, 261-269, 1997.

Lia H. Campbell, Keith T. Borg and Salvatore J. Arrigo. Differential Effects of Intronic and Exonic Locations of the Human Immunodeficiency Virus Type 1 Rev-Responsive Element. Virology, 219, 423-431, 1996.

Lia H. Campbell, Keith T. Borg, Julia K. Haines, Randall T. Moon, Daniel R. Schoenberg and Salvatore J. Arrigo. Human Immunodeficiency Virus Type 1 Rev Is Required In Vivo for Binding of Poly(A)- Binding to Rev-Dependent RNAs. Journal of Virology, 68:9, 5433-5438, 1994.